END

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

RECURSIVE NONLINEAR FILTERING THROUGH
COMPLETE LINEARIZATION

T. Dyson

G.R.D., Incorporated
Warminster Operations
Warminster, PA   18974

April 1985

Prepared for:

DTIC
ELECTE
JUN 1 7 1985
S      D
G

85   5 20 029

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>1 | 2. GOVT ACCESSION NO.<br>AD-A155 029 | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br><br>Recursive Nonlinear Filtering through<br>Complete Linearization | | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report<br>March 1984 - March 1985 |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br><br>Timothy F. Dyson | | 8. CONTRACT OR GRANT NUMBER(s)<br><br>N00014-84-C-0233 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br><br>GRD, Inc.<br>444 Jacksonville Road<br>Warminster, PA 18974 | | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br><br>NR 277-322/410 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Mathematical Sciences Division<br>Office of Naval Research<br>Arlington, VA 22217 | | 12. REPORT DATE<br>April 1985 |
| | | 13. NUMBER OF PAGES<br>76 |
| 14. MONITORING AGENCY NAME & ADDRESS(If different from Controlling Office)<br>Naval Air Systems Command<br>NAVAIR Code 340R<br>Washington, D.C. 20361 | | 15. SECURITY CLASS. (of this report)<br><br>Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release; Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Also presented in part at 18th Princeton Information Sciences and Systems
Conference and 24th IEEE CDC. To be submitted as part of a Ph.D. Thesis
of T.F. Dyson to the EECS Dept., Princeton University.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Filtering, Differential Geometry, Linearization, Transformation
Theory, Extended Kalman Filters

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

ABSTRACT

Recursive filtering forms the foundation of all tracking and position
estimation tasks when the target undergoes dynamic motion relative to the
observer. Moreover, recursive filtering has an important role in stochastic
signal detection. The typical application almost invariably involves dynamics
equations which depend nonlinearly on the physical variables and sometimes
even the sensor (observation) equation involves a nonlinear relation to the

DD FORM<br>1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

quantities to be estimated. In these common situations the practitioner nearly always resorts to a form of approximate linearization of the describing equations to allow the use of the now classical Kalman-Bucy filter equations. Although highly successful in many useful application, it is commonly recognized that a variety of problems simply will not satisfactorily succumb to the usual "Extended Kalman" methods. For these situations present research in nonlinear filtering theory hold major portent. Indeed, the new understanding of nonlinear problems developed over the past 10 years leads to optimism for better, more refined approximate filtering techniques which may prove to be both accurate and more widely applicable.

Alternatively., dynamics equations which naturally formulated are non-linear may sometimes be maneuvered into a form which is exactly linear. This phenomenon has been well-known in deterministic problems for a long time. That the same situations might arise in a stochastic setting seemed to be obvious but was never characterized in detail except for the simplest of situations. In this report such system state variable equations are completely characterized and a method for converting them to an equivalent linear representation given. In addition very simple tests are developed to determine when a nonlinear stochastic differential equation of the diffusion type can be completely linearized by these methods.

Along the way it is pointed out that, although highly improbable in real-world situations, under the right conditions, certain nonlinear filtering problems become completely linear ones by these tests and methods and hence yield to the well-known, exact Kalman-Bucy filtering. Beyond these trivial and rather contrived situations, the existence of linearizable equations of "motion" suggest potentially more accurate approximate filters based on the transformed state variables. One such situation is investigated here and it is shown empirically that lower estimation error is obtained from this technique over an alternative extended Kalman filter for the same real-world problem.

Accession For

NTIS GRA&I

DTIC TAB

Unannounced

Justification

By

Distribution/

Availability Codes

Dist | Avail and/or Special

A/

# RECURSIVE NONLINEAR FILTERING THROUGH

## COMPLETE LINEARIZATION

Timothy F. Dyson
444 Jacksonville Road
Warminster, PA 18974

## ABSTRACT

Recursive filtering forms the foundation of all tracking and position estimation tasks when the target undergoes dynamic motion relative to the observer. Moreover, recursive filtering has an important role in stochastic signal detection. The typical application almost invariably involves dynamics equations which depend nonlinearly on the physical variables and sometimes even the sensor (observation) equation involves a nonlinear relation to the quantities to be estimated. In these common situations the practitioner nearly always resorts to a form of approximate linearization of the describing equations to allow the use of the now classical Kalman-Bucy filter equations. Although highly successful in many useful applications, it is commonly recognized that a variety of problems simply will not satisfactorily succumb to the usual "Extended Kalman" methods. For these situations, current research in nonlinear filtering theory hold major portent. Indeed, the new understanding of nonlinear problems developed over the past 10 years leads to optimism for better, more refined approximate filtering techniques which may prove to be both accurate and more widely applicable.

Alternatively, dynamics equations, which naturally formulated are nonlinear, may sometimes be maneuvered into a form which is exactly linear. This phenomenon has been well-known in deterministic problems for a long time. That the same situations might arise in a stochastic setting seemed

to be obvious but was never characterized in detail except for the simplest of situations. In this report such system state variable equations are completely characterized and a method for converting them to an equivalent linear realization given. In addition, very simple tests are developed to determine when a nonlinear stochastic differential equation of the diffusion type can be completely linearized by these methods.

Along the way it is pointed out that, under the right conditions, although highly improbable in real-world situations, certain nonlinear filtering problems become completely linear ones by these tests and methods and hence yield to the well-known, exact Kalman-Bucy filtering. Beyond these trivial and rather contrived situations, the existence of linearizable equations of "motion" suggest potentially more accurate approximate filters based on the transformed state variables. One such situation is investigated here and it is shown empirically that lower estimation error is obtained from this technique over an alternative extended Kalman filter for the same real-world problem.

## TABLE OF CONTENTS

TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# TABLE OF CONTENTS

| TITLE | PAGE |
|---|---|

### LIST OF TABLES (Continued)

# CHAPTER 1

## INTRODUCTION

This report addresses the problem of sample path estimation of partially observed stochastic processes. Also referred to as filtering, this problem arises in diverse applications: communication, economics, radar and sonar. Specifically, this report considers an approximate filtering approach for problems in which the state variables are governed by certain nonlinear stochastic differential equations which are closely related to a linear stochastic differential equation. The suggested approximation scheme is based on developing extended Kalman filters from the related linear representation to avoid the sometimes highly nonlinear functionals occurring in the nonlinear state equations.

Although completely linearizable deterministic nonlinear differential equations are well-known [7] and completely characterized [8,9], and equivalent linearization ideas for stochastic nonlinear differential equations surmised for some time [10,11] it has remained to specifically describe and characterize in detail the conditions under which this transformation from nonlinear representation to linear representation would occur in the stochastic setting. Chapter 2 accomplishes this detailed characterization for all time invariant nonlinear stochastic differential equations. Chapter 3 studies an approximate filtering application using the equivalent linear representation ideas in an amplitude and phase tracking problem.

## The Filtering Problem

The general filtering problem, a specialization of which is dealt with in this report, is expressed in the following way:

Let $(\Omega, F, P)$ be a fixed underlying probability space on which there is an increasing family of $\sigma$-fields, $\{F_t, 0 \le t \le T\}$. The **signal** or **state process**, $\{x_t, t \ge 0\}$, is $F_t$ measurable for each $t \in [0, T]$ and is assumed to be a homogeneous Markov process and a diffusion process satisfying a stochastic differential equation of the form,

$$dx_t = f(x_t)dt + g(x_t)dw_t \tag{1.1}$$

where, $x_t \in \mathbf{R}^n$; $w_t \in \mathbf{R}^m$ is a vector valued standard Brownian motion process defined on $(\Omega, F, P)$. The process $\{x_t, t \ge 0\}$ has generator $L$, where

$$L\phi(x) = \{\sum_1^n f_i(x) \frac{\partial}{\partial x_i}\} + \frac{1}{2} \sum_{i,j=1}^n \{\sum_{k=1}^m g_{ik}(x)g_{jk}(x)\}\frac{\partial^2}{\partial x_i \partial x_j}\} \phi(x)$$

for all $\phi$ in the domain of $L$. The functions f, g are assumed to be measurable mappings and to satisfy continuity and growth conditions (e.g., [1,2]) sufficient to insure that $\{x_t\}$ exists.

In the filtering problem $\{x_t\}$ is not directly observed, but is observed through a related process $\{y_t, t \ge 0\}$ satisfying

$$dy_t = h(x_t) dt + dv_t \tag{1.2}$$

where $\{v_t, t \ge 0\}$ is a standard Brownian motion, $F_t$-measurable and $h(\cdot)$ is some mapping from $\mathbf{R}^n$ to range of $\{y_t\}$, say $\mathbf{R}^k$. The problem is then to compute conditional statistics of $\{x_t\}$ given only the $\sigma$-field generated by $\{y_s, 0 \le s \le t\}$.

## Present Status of Filtering Theory

Considerable progress has been made in this problem by dealing with the so-called Duncan-Mortensen-Zakai equation [2,11,12] for the evolution of the unnormalized conditional density, $\rho(x, t)$, given by

$$\rho(x,t) = \rho(x_o, 0) + \int_o^t L^*\rho(x,s)ds + \int_o^t \rho(x,s)h(x)dy_s \tag{1.3}$$

where $L*$ is the formal adjoint of $L$ and the last integral is to be interpreted as a stochastic integral [1,2]. The progress in understanding the nature of nonlinear filtering is due primarily to the bilinear form of this equation and the connections made between (1.3), the theory of nonlinear partial differential equations and nonlinear systems theory [4,5,6].

In particular, connections between the Lie Algebraic properties of the algebra of operators induced by (1.3), or more precisely the Stratonovich form for (1.3),

$$L = \{L* - \frac{1}{2} h^2, h\}_{LA}$$

and the finiteness conjecture of Brocket [4] have proven to be the most fruitful avenues. Although the fundamental Homomorphism Conjecture [4,6,25] required to close the loop on nonlinear filtering theory has been elusive in the general case, (see however, [19,21,25 & 26) many striking results have evolved from this line of reasoning, for example [13-20].

The status of approximate filtering for this problem framework is less satisfactory. In a limited class of problems the aforementioned Lie Algebraic studies have produced reasonable results [13,14,21]. More generally applicable approaches include the time worn extended Kalman filter and second order approximations [24] and a few less well known series expansion techniques [21-23]. However, it has been strongly suggested in the literature that Lie Algebraic techniques may be a profitable vein of inquiry for approximate filtering in a more general context than has so far been attempted [6,12a,13,25, especially 27, 28].

## Summary of Report

The remainder of this report deals with a rather different tact for approximate filtering applications. The approach taken is to consider problems in which the state process, in addition to the conditions above, admits pointwise transformation to a potentially more useful representation. Chapter 2 develops the details of the transformation theory for a linear representation. Although this material is built straight forwardly upon work in deterministic linearization [7,9], very concise, as well as constructive, procedures are developed to

a. test any given nonlinear stochastic differential state equation for the extistance of an alternate "controllable" linear representation, and

b. exactly determine the smooth transformation of state variables providing this alternate representation.

Chapter 3 then indicates, by way of an interesting example of practical significance, how this transformation theory can be applied to formulate approximate state estimation filters. The example studied is subjected to Monte Carlo simulation to demonstrate the potential performance advantage of the suggested a⊦. ⋅ion of the transformation approach over other extended Kalman filters.

In an analogous fashion we have the vector spaces

$$\Lambda^2(\mathbf{R}^n) = \{\omega(x) \mid \omega(x) = \sum_{i,j=1}^{n} \omega^{ij}(x)dx^i \wedge dx^j, \ \omega^{ij}(c)C^\infty\},$$

$\Lambda^3(\mathbf{R}^n), \cdots, \Lambda^n(\mathbf{R}^n)$ of p-forms (p=2,3,...,n). We also identify $\Lambda^0(\mathbf{R}^n)$ with $F(\mathbf{R}^n)$. The collection of all p-forms over $\mathbf{R}^n$ ($o \leq p \leq n$) together with the exterior product defined as above, and extended in the obvious way to any $\omega \in \Lambda(\mathbf{R}^n)$, defines the Exterior Algebra, $(\Lambda(\mathbf{R}^n), \wedge, F(\mathbf{R}^n))$, or simply $\Lambda(\mathbf{R}^n)$.

If in addition we have an inner product on $\mathbf{R}^n$ then using (2.14) let $x \in \Lambda^{p\star}(\mathbf{R}^n)$ and $y \in \Lambda^\gamma(\mathbf{R}^n)$ we have a pairing between $\Lambda(\mathbf{R}^n)$ and its dual:

$$\langle x \mid y \rangle = \begin{cases} 0 & , \ p \neq \gamma \\ \det(x_i, y^j), & p = \gamma \end{cases} \tag{2.15}$$

where $\Lambda^p(\mathbf{R}^n)$ = space of exterior p-forms over $\mathbf{R}^n$.

**We summarize the major properties of p-forms:**

(a)   if dim $T(\mathbf{R}^n)$ = n then $\Lambda^n(\mathbf{R}^n)$ is 1-dimensional

(b)   $\Lambda^p(\mathbf{R}^n)$ has dimension $\binom{n}{p}$

(c)   a basis for $\Lambda^p(\mathbf{R}^n)$ is $dx^{i_1} \wedge \cdots \wedge dx^{i_p}$; $i_1, \cdots, i_p = 1, \cdots, n$

(d)   $dx^m \wedge dx^m = 0$; if $\omega \in \Lambda(\mathbf{R}^n)$, $\omega \wedge \omega = 0$

(e)   (Exterior Differentiation) $\underline{d}: \Lambda^p(\mathbf{R}^n) \to \Lambda^{p+1}(\mathbf{R}^n)$; $p \leq$ dim $T(\mathbf{R}^n)-1$

   If $\omega = \sum_{1}^{n} a_{i_1}, \cdots, i_p(x)dx^{i_1} \wedge \cdots \wedge dx^{i_p}$

   then $\underline{d}\omega = \sum_{j=1}^{n} \sum_{1}^{n} \frac{\partial a_{i_1 \cdots i_p}}{\partial x_j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}$

(f)   $\Lambda^0(\mathbf{R}^n) = F(\mathbf{R}^n)$

(g)   by (a), if $\omega \in \Lambda^n(\mathbf{R}^n)$ and dim $T(\mathbf{R}^n)$ = n, $\underline{d}\omega = 0$

(h)   If $\underline{d}\omega = 0$ then $\omega$ is Closed,

The natural pairing between $T(\mathbf{R}^n)$ and $\Lambda^1(\mathbf{R}^n)$ is expressed by

$$\langle \tau, \eta \rangle_x = \sum_1^n \sum_1^n \tau^i(x)\, \eta_j(x) \left\langle dx^j, \frac{\partial}{\partial x_i} \right\rangle_x$$

$$= \sum_1^n \tau^i(x)\, \eta_i(x)$$

(2.14)

where the dual basis elements satisfy the relation

$$\left\langle dx^j, \frac{\partial}{\partial x_i} \right\rangle_x = \delta_{ij}$$

Hence $\langle \cdot, \cdot \rangle : T(\mathbf{R}^n) \times \Lambda^1(\mathbf{R}^n) \to F(\mathbf{R}^n)$ is a bilinear mapping of the tangent space and its dual, and thus defines the action of $\eta \in \Lambda^1(\mathbf{R}^n)$ as a linear functional on $T(\mathbf{R}^n)$ or vice versus.

From the vector space structure on $\Lambda^1(\mathbf{R}^n)$ it is possible to construct an algebra of differential forms known as the Exterior Algebra, $\Lambda(\mathbf{R}^n)$. The product operation used to construct $\Lambda(\mathbf{R}^n)$ is called the "wedge" or exterior product. It is a skew symmetric multiplication of differential forms defined, in coordinates, as follows:

For $\nu, \eta \in \Lambda^1(\mathbf{R}^n)$, $\nu(x) = \sum_1^n \nu_i(x) dx^i$ and $\eta(x) = \sum_1^n \eta_i(x) dx^i$

$$\nu \wedge \eta(x) = \sum_{i=1}^n \sum_{j=1}^n \nu_i(x)\eta_j(x) dx^i \wedge dx^j$$

where

$$dx^i \wedge dx^j = \begin{cases} 0 & i = j \\ -dx^j \wedge dx^i & i \neq j \end{cases}$$

## 3. SINGLE INPUT N-DIMENSIONAL STOCHASTIC DIFFERENTIAL EQUATIONS

### Geometric Preliminaries

This section addresses the n-dimensional problem. Because of the added complexity of this problem, we first review several well known results from Differential Geometry and the Exterior Calculus which will be useful in what follows. Everything we will do is in $\mathbf{R}^n$, so the discussion is restricted to these spaces. Much of the general material presented comes from [10] and [11].

Smooth ($C^\infty$) mappings from $\mathbf{R} \to \mathbf{R}^n$ induce vector fields over $\mathbf{R}^n$. The natural (Euclidean) coordinate system on $\mathbf{R}^n$, $(x_1, \dots x_n)$, induces an atlas on $\mathbf{R}^n$. With this atlas, $\mathbf{R}^n$ is viewed as a differentiable manifold and the induced vector fields over $\mathbf{R}^n$ can be identified with the vector space of linear functionals on the space of functions on this manifold, $F(\mathbf{R}^n): \mathbf{R}^n \to \mathbf{R}$. Under this identification, a natural basis (in a coordinate neighborhood) on the space of vector fields over $\mathbf{R}^n$ (i.e. the Tangent Space to $\mathbf{R}^n - T(\mathbf{R}^n)$) is

$$\{\frac{\partial}{\partial x_1}, \dots \frac{\partial}{\partial x_n}\} \quad .$$

Thus a "tangent" vector field in $T(\mathbf{R}^n)$ is given (locally) by

$$\tau(x) = \sum_1^n \tau^i(x) \frac{\partial}{\partial x_i} \quad , \tag{2.13}$$

where $x \in \mathbf{R}^n$ and the $\tau^i(\cdot)$ are $C^\infty$ functions on $\mathbf{R}^n$. The dual space to the tangent space $T(\mathbf{R}^n)$, is $T^*(\mathbf{R}^n) \sim T(\mathbf{R}^{n*}) \underline{\underline{\Delta}} \Lambda^1(\mathbf{R}^n)$ and carries the corresponding dual basis $\{dx^1, \dots, dx^n\}$. Elements of $\Lambda^1(\mathbf{R}^n)$ are called 1-forms (differential forms of degree 1) and have the coordinate respresentation

$$\eta(x) = \sum_1^n \eta_i(x) dx^i.$$

h(x) is at least once continuously differentiable. Then applying extended

Kalman theory (EKT) to the transformed equations gives the state estimate

$$d\bar{z}_t = a\bar{z}_t dt + P_t \bar{H}_t (dy_t - h_c(\bar{z}_t)dt)$$

$$\bar{H}_t = \frac{\partial h_c}{\partial z}(\bar{z}_t)$$  (2.12)

$$\dot{P}_t = 2aP_t + b^2 - P_t^2 \bar{H}_t^2$$

The approximation enters only through the function $h_c$ in this example.

Based on this, one might expect $\bar{z}_t$ to be a better approximation to the

conditional mean estimate $E\{z_t | F_y^t\}$ than the corresponding approximation for

$E\{x_t | F_y^t\}$ developed through application of EKT to the original stochastic

differential equations in certain cases. This conjecture is investigated in

Chapter 3 for an interesting practical example.

Example 3  Consider the s.d.e.

$$dx_t = \frac{x_t q_-(x_t)}{q_+(x_t)} \left(a - b^2 \frac{(4 + q_-^2(x_t))q(x_t)}{q_+^2(x_t)}\right)dt + b \frac{q_-^2(x_t)}{q_+(x_t)} dw_t$$

where $q_\pm(x) = (1 \pm x^2)$. With the initial state value $|x_0| < 1$

with probability 1, the diffusion process $\{x_t, t \geq 0\}$ has as its state

space the open interval $(-1,1)$. Applying Theorem 1, it is found that under

$$\phi(x) = \frac{x}{1-x^2} = u$$

$$du_t = au_t dt + bdw_t .$$

The next section makes use of a 2-dimensional generalization of this

example.

for which the equivalent Gauss-Markov diffusion would be

$$dz_t = az_t \, dt + bdw_t \qquad (2.10)$$

If in addition the observation is of the form

$$dy_t = cx_t(1 + \tfrac{1}{3} x_t^2)dt + dv_t \qquad (2.11)$$

and $z_o = x_o(1 + \tfrac{1}{3}x_o^2)$ is distributed $G(0, \sigma_o^2)$ then the optimal filtering

problem for $\{x_t, \ t \geq 0\}$ is solved by the Kalman-Bucy equations for the equivalent

problem (2.10) and (2.11). The conditional density is

$$p(x,t|y_s, \ 0 \leq s \leq t) = \frac{(1 + x^2)}{(2\pi P(t))^{1/2}} \exp \ \{- \tfrac{1}{2}[x(1 + \tfrac{1}{3} x \ )- \hat{z}_t]^2/P(t)\}$$

where $\{\hat{z}_t, \ P(t)\}$ are the solution of the Kalman filter for (2.10) and (2.11).

NOTE: The unappealing Gaussian initial condition is easily circumvented.

The work of Benes and Karatzas [9a] or Makowski [9b] can be applied to handle

arbitrary initial state distributions.

The foregoing example is rather contrived; it seems unlikely in practice

that one would encounter a problem which is both transformable to a linear

s.d.e. and has just the right observation drift. However, in the next example

the potential application of the transformation theory to develop approximate

solutions is indicated. Further justification for why one might use this

approach is postponed until Chapter 3.

Example 2  Assume the state equation satisfies the conditions of Theorem 1 and

has the equivalent linear representation under transformation, $z_t = \phi(x_t)$

$$dz_t = az_t dt + bdw_t \quad .$$

Assume the observation equation is

$$dy_t = h(x_t)dt + dv_t$$

where $h \neq \phi$ and define $h_c = h \circ \phi^{-1}$.  Assume also that

In this case the $C^\infty$ transformation relating the two diffusions is

$$z = \phi(x) = \frac{b}{a} \frac{Q(x)}{g(x)} .$$

Remarks:

The theorem as stated is valid only for $a \neq 0$ and thus restricts $S_x = \mathbb{R}$.

Generalizing slightly by using $a\phi(x)+c$ on the r.h.s. of (2.7) and allowing $a=0$ admits situations like $S_x = \mathbb{R}^+ - \{0\}$. The only other modification required for this case is that $\phi(x)$ is constructed by solving the latter half of (2.6) instead of (2.7). This, for example, covers the well known case of a bilinear s.d.e.

We give several examples of the use of Theorem 1.

<u>Example 1</u>  Consider the diffusion

$$dx = \left[ \frac{ax_t(1 + \frac{1}{3}x_t^2)}{(1 + x_t^2)} - \frac{b^2 x_t}{(1 + x_t^2)^3} \right] dt + \frac{b}{(1 + x_t^2)} dw_t \qquad (2.8)$$

Here,

$$Q(x) = ax \frac{(1 + \frac{1}{3} x^2)}{(1 + x^2)}$$

$$g(x) = \frac{b}{(1 + x^2)}$$

so

$$Q(x)g'(x) - g(x)Q'(x) + \alpha g(x)$$

$$= \frac{\alpha b - ab}{(1 + x^2)}$$

By choosing $\alpha = a$, Theorem 1 is satisfied and the transformation is given by

$$z = \phi(x) = x(1 + \frac{1}{3} x^2) \qquad (2.9)$$

The only consistent solution for this equation occurs when the bracketed term is identically zero. Substituting back into the remaining equation gives

$$\frac{\partial \phi(x)}{\partial x} \left[ f(x) - \frac{1}{2} \frac{\partial g(x)}{\partial x} g(x) \right] = a\phi(x) \tag{2.7}$$

Define $Q(x) = f(x) - \frac{1}{2} \frac{\partial g}{\partial x} g(x)$ and notice that

$$\frac{\partial^2 \phi}{\partial x^2}(x) Q(x) + \frac{\partial \phi}{\partial x}(x) \frac{\partial Q}{\partial x}(x) = a \frac{\partial \phi}{\partial x}(x).$$

Using these it follows that

$$\frac{\partial \phi}{\partial x}(x) \left\{ Q(x) \frac{\partial g}{\partial x}(x) - \frac{\partial Q}{\partial x}(x) g(x) \right\} = - a \frac{\partial \phi}{\partial x}(x) g(x).$$

Then letting $[Q, g](x) = Q(x) \frac{\partial g}{\partial x}(x) - \frac{\partial Q}{\partial x}(x) g(x)$,

and combining with (6) we have the necessary and sufficient equation,

$$\frac{\partial \phi}{\partial x}(x) \{ [Q,g](x) + a g(x) \} = 0 \quad .$$

The desired transformation may be found by solving (2.7) and using (2.6):

$$\phi(x) = \frac{1}{a} Q(x) \frac{\partial \phi(x)}{\partial x} = \frac{b}{a} \frac{Q(x)}{g(x)} \quad .$$

The equivalent linear s.d.e. is

$$dz_t = d\phi(x_t) = a\phi(x_t)dt + bdw_t$$

$$= az_t dt + bdw_t$$

Summarizing, we have:

<u>THEOREM 1</u> The $C^\infty$ non-degenerate diffusion on $\mathbb{R}$, specified by the s.d.e.

$dx_t = f(x_t)dt + g(x_t)dw_t$, is equivalent to a linear s.d.e. on $\mathbb{R}$ of the form

$$dz_t = az_t dt + bdw_t \quad , \quad (a \neq 0)$$

if and only if

(i) $\quad |g(x)| > 0$

(ii) $\quad Q(x) = f(x) - \frac{1}{2} \frac{\partial g(x)}{\partial x} g(x), \; Q(0) = 0$

(iii) $\quad Q(x) \frac{\partial g(x)}{\partial x} - g(x) \frac{\partial Q(x)}{\partial x} + a g(x) = 0 \quad \forall \, x \, \varepsilon \, \mathbb{R}$

## 2. SCALAR DIFFUSIONS

The general scalar diffusion model of interest is

$$dx_t = f(x_t)dt + g(x_t)dw_t. \tag{2.3}$$

Here $\{w_t, \; t \geq 0\}$ is a standard Brownian Motion on $\mathbb{R}$ and the functions $f(x)$, $g(x)$ satisfy the conditions stated in Section 1. The state space, $S_x$, is also assumed to be an open set in $\mathbb{R}$. The reason for this restriction is to ensure that $(f(x), g(x))$, $x \in$ interior $(S_x)$ completely determine the characteristic operator of the diffusion $\{x_t, t \geq 0\}$, $\overset{\circ}{A}$, at every point in $S_x$. In simpler terms, this assumption excludes all but natural boundaries from the state space. More will be said about this in the remarks to follow.

What is sought is a smooth mapping $\phi$, taking $x_t$ into $z_t$ in such a way that the s.d.e. corresponding to 2.3 is linear and of the form

$$dz_t = az_t \, dt + bdw_t \tag{2.4}$$

where $a,b$ are constants and $a \neq 0$. Assuming that the transformation is at least twice continuously differentiable, the Ito differential rule gives

$$d\phi(x_t) = \left[ \frac{\partial \phi(x_t)}{\partial x} f(x_t) + \frac{1}{2} \frac{\partial^2 \phi(x_t)}{\partial x^2} g^2(x_t) \right] dt + \frac{\partial \phi(x)}{\partial x} g(x_t)dw_t \tag{2.5}$$

Letting $z_t = \phi(x_t)$ and equating corresponding terms in (2.4) and (2.5) yields

$$\frac{\partial \phi(x)}{\partial x} f(x) + \frac{1}{2} \frac{\partial^2 \phi(x)}{\partial x^2} g^2(x) = a\phi(x)$$

$$\frac{\partial \phi(x)}{\partial x} g(x) = b \tag{2.6}$$

for all $x \in S_x$.

To solve these simultaneous equations, first note that due to the second part of (6),

$$\frac{\partial}{\partial x} \left[ \frac{\partial \phi(x)}{\partial x} g(x) \right] = \frac{\partial^2 \phi(x)}{\partial x^2} g(x) + \frac{\partial \phi(x)}{\partial x} \frac{\partial g(x)}{\partial x} = 0 \; .$$

5 discusses two further aspects of the transformation idea. The first is the targeting of nonlinear s.d.e.'s in the transformation by way of an example using the Benes state equation. The second topic deals with mapping nonlinear s.d.e.'s to higher state dimension linear representations within the context of the material in Section 3. The final section summarizes the main results of the chapter.

to a linear stochastic differential equation. In so doing, this work is a
direct extension of the ideas, but more importantly a simplification of the
methods, in [12,22] to nonlinear stochastic differential equations.

In the remainder of this chapter, the class of problems defined by
nonlinear s.d.e.s which can be pointwise transformed into an equivalent
version in the class of linear s.d.e.s is characterized in detail. A proce-
dure for constructing the required transformation and for recognizing when a
particular nonlinear s.d.e. is amenable to this transformation are determined.

In the next section, the 1-dimensional problem is discussed to delineate
the background ideas involved in the more difficult n-dimensional cases. The
existence of an equivalent linear representation for a scalar nonlinear
s.d.e. is characterized in terms of a certain ordinary differential equation.
This easy and straightforward case is presented first because of its expository
value for the uninitiated reader. Further, this discussion allows a convenient
introduction of both an (admittedly contrived) optimal filtering application
and a potentially useful approximate filtering application of the general
transformation theory. Section 3 addresses the transformation of n-dimensional
nonlinear s.d.e.'s. Tools from Geometry, which are used in the n-dimensional
setting, are summarized first. These tools are employed to describe
necessary and sufficient conditions (similar to those in [12]) a nonlinear
s.d.e. must satisfy to admit transformation to a linear s.d.e. target
representation. Simplified practical tests are devised for determining
existance of the necessary transformation of state coordinates, and a simple
constructive method is provided for determining this transformation. Examples
are also given. Section 4 presents two straight forward extensions of the
results of previous sections to multi-input/multi-output and time varying
stochastic differential equations. Examples of each are discussed. Section

The standard problem in nonlinear filtering is: given (2.1) and an observed process $\{y_t, t \geq 0\}$, related to $\{x_t, t \geq 0\}$ by the s.d.e.

$$dy_t = h(x_t)dt + dv_t \ , \tag{2.2}$$

where $\{y_t, t \geq 0\} \in \mathbb{R}^d$, $h: \mathbb{R}^n \to \mathbb{R}^d$ is Borel measurable and $\{v_t, t > 0\}$ is vector valued Brownian Motion, construct an estimate of $\{x_t, t \geq 0\}$ given $\{y_s, 0 \geq s \geq t\}$. Optimal filtering solutions, in the mean square error sense, have been constructed in a few cases: Gauss-Markov case (Kalman-Bucy theory); the Benes problem, where $f(x)$ satisfies a Riccati differential equation, $g(x)$ is a constant and $h(x)$ is linear (not affine) in x [2,3]; conditionally Guassian processes [4]; discrete state Markov problems [5]; and the multidimensional cases constucted by W. Wong [6], to name several. Approximate filtering solutions for other situations are somewhat limited and include the extended Kalman techniques and a handful of series expansion (or related) approaches when "small" nonlinearities are involved (see e.g., [19], [20] and [21]).

Brockett [7], Baras [9] and others [22] have studied the idea of equivalence classes of filtering problems by considering guage transformations, pointwise state transformations, and what Baras has called Group Invariance methods. The concept here is to identify when a given problem, equations (2.1) and (2.2), or (2.1) alone, is merely a disguised version of another perhaps better understood problem. Similar notions have been discussed in the deterministic control systems literature as far back as 1973 by Krener [22] and more recently by Su [12], and Hunt, Su and Meyer [13].

Motivated by these equivalence ideas and by a desire to obtain alternatives to standard extended Kalman filters, this report develops a set of constructive tests through which it can be determined when a diffusion satisfying the s.d.e. (2.1) can be transformed pointwise to another, possibly more useful form. In particular, the bulk of the report deals with transformation

# CHAPTER 2

## TRANSFORMATION THEORY FOR NONLINEAR STOCHASTIC

## DIFFERENTIAL EQUATIONS ON $\mathbf{R}^n$

### 1. INTRODUCTION

Nonlinear filtering problems frequently model the "signal" as a diffusion process satisfying a stochastic differential equation (s.d.e.) of the form

$$dx_t = f(x_t)dt + g(x_t)dw_t , \qquad (2.1)$$

where $\{w_t, t \geq 0\}$ is a standard Brownian Motion on $\mathbf{R}^m$, $m \geq 1$, and $\{x_t, t \geq 0\}$ $\epsilon \; \mathbf{R}^n, n \geq m$. $g(x) = (g^i_j(x))$ is a Borel measurable (in general matrix valued) function $(x) \epsilon \; \mathbf{R}^n \rightarrow \mathbf{R}^n \otimes \mathbf{R}^m$ and $f(x) = (f_i(x))$ is a Borel measurable (vector valued) function $(x) \epsilon \; \mathbf{R}^n \rightarrow \mathbf{R}^n$. In this case, the s.d.e. (2.1) is said to be of Markovian type [1]. A s.d.e. thus generalizes the notion of an ordinary differential equation by adding the effects of random fluctuation.

For the purposes of this report, the drift $(f(x))$ and diffusion $(g(x))$ coefficients, and the process defined by (2.1) are restricted further:

    (i) $f,g$ are assumed to be continuously differentiable $(C^\infty)$ functions of their arguments.

    (ii) we specifically exclude those s.d.e.'s in which $g(x) = 0$ for some $x$ in the state space, $S_x$.

The first restriction ensures that the needed differentiability of $f$ and $g$ is always available and allows one to avoid the tedium of keeping track of the degree of differentiation. The second restriction voids degenerate stochastic differential equations and thereby guarantees that the equations we wish to solve (see below) do not exhibit explosions. Diffusion processes governed by s.d.e.'s satisfying these conditions will henceforth be referred to as $C^\infty$ non-degenerate diffusions.

[15] V.E. Benes, Exact finite dimensional filters for certain diffusions with nonlinear drift, Stochastics, 5, 1981, pp 65-92.

[16] W.S. Wong, New classes of finite dimensional filters, Syst. and Cont. Lett, 3, 1983, pp. 155-164.

[17] D. Ocone, J. Baras and S.I. Marcus, Explicit filters for diffusions with certain nonlinear drifts, Stochastics, 8, 1982, pp 1-16.

[18] D. Ocone, Finite ·dimensionally computable statistics and estimation algebras in nonlinear filtering, Proc. 1981 International Symposium on the Mathematical Theory of Networks and Systems, Santa Barbara, CA, 1981.

[19] M. Chaleyat-Maurel and D. Michel, Un theoreme de non existence de filtre de dimension finie, C.R. Acad. Sc. Paris, t. 296, 20 June 1983, Serie I-933.

[20] _____, Une condition suffisante de non-existence d'un filtre de dimension finie, C.R. Acad, Sc. Paris, t. 297, 26 September 1983, Serie I-203.

[21] H. Sussmann, Approximate finite dimensional filters for some nonlinear problems, Stochastics, 7, 1982, pp 183-203.

[22] W. Hopkins, Multiplicative asymptotic expansion for filtering linear diffusions with small polynominal nonlinearities, IEEE Conf. on Decision and Control, 1982, pp 195-200.

[23] H. Sorenson, A Bayesian Approach to Nonlinear filtering and plant identification, preprint, May 1982.

[24] A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

[25] M. Hazewinkel, The Linear Systems Lie Algebra, the Segal-Shale-Weyl representation and all Kalman-Bucy filters, preprint, Econometric Institute R#8130/M, Erasmus University, 1982.

[26] M. Hazewinkel, S.I. Marcus & H.J. Sussmann, Non existence of finite dimensional filters for conditional statistics of the cubic sensor problem, preprint, Stichting Mathematishch Centrum, 1983.

[27] M. Hzewinkel and S.I. Marcus, Some Results and Speculations on the Role of Lie Algebras in Filtering, in Stockastic Systems: The Mathematics of Filtering and Identification and Applications, 591-604, 1981.

[28] B. Hanzon and M. Hazewinkel, On Identification of Linear Systems and the Estimation Lie Algebra of the Associated Nonlinear Filtering Problem, preprint Econometric Institute Report 8122/M, Erasmus University, 1981.

Chapter 1 - References

[1] E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York 1971.

[2] R.S. Lipster and A.N. Shirgaev, Statistics of Random Processes I, Springer-Verlag, New York, 1977.

[3] M.N.A. Davis and S.I. Marcus, An Introduction to nonlinear filtering, Stochastic Systems: The Mathematics of Filtering and Identification and Applications, M. Hazeninkel and J.C. Williams, eds., Reidel, Dordiecht, The Netherlands, 1981, pp 565-572.

[4] R.W. Brockett, Remarks on finite dimensional nonlinear estimation, Analysis des systemes, C. Labry, ed., Bordeaux, 1978.

[5] R.W. Brockett and J.M.C. Clark, The geometry of the conditional density equation, Analysis and Optimization of Stochastic Systems, O.L.R. Jacobs et.al., eds., New York, 1980, pp 299-309.

[6] S.I. Marcus, Algebraic and Geometric Methods in nonlinear filtering, SIAM J. Control and Optimization, 22, No. 6, November 1984, pp 817-844.

[7] A.J. Krener, on the Equivalence of Control Systems and the Linearization of Nonlinear Systems, SIAM J. Control, 11, No. 4, November 1973.

[8] V. Respondek, preprint, Geometric Methods in Linearization of Control Systems, 19

[9] R.Su, On the Lienar Equivalent of Nonlinear Systems, Syst. and Control Letters, 2, No. 1, July 1982 pp 48-52.

[10] R.W. Brockett, Nonlinear Systems and differential geometry, Proc. IEEE, 64 (1976) pp. 61-72.

[11] _____, Classification and equivalence in estimation theory, Proc. 18th IEEE Conference on Decision and Control, Ft. Lauderdale, FL, 1979, pp. 172-175.

[12] _____, Nonlinear Systems and Nonlinear Estimation Theory, Stochastic Systems. The Mathematics of Filtering and Identification and Applications, M. Hazewinkel and J.C. Williams, eds., Reidel, Dorieckt, The Netherlands, 1981, pp 441-478.

[12a] M. Hazewinkel, On deformations, approximations and nonlinear filtering, Syst. & Cont. Lett. 1, No. 1, July 1981, pp 32-36.

[13] M. Hazewinkel and S.I. Marcus, On Lie Algebras and finite dimensional filtering, Stochastics, 7, 1982, pp 29-62.

[14] G.L. Blankenship, C.H. Liu and S.I. Marcus, Asymptotic Expansions and Lie Algebras for Some Nonlinear Filtering Problems, IEEE Transactions on Automatic Control, AC-28, No. 7, 1983, pp 787-797.

If $\omega = \underline{d}\sigma$ then $\omega$ is Exact

(i)  Poincaré's Lemma:  $\underline{d}(\underline{d}\omega) = 0$, $\omega \epsilon \Lambda^P(\mathbf{R}^n) \forall\ p \leq \dim\ T(\mathbf{R}^n)$

## Linear Equivalence Transformations in n-Dimensions

In the spirit of Section II, and using the tools of the previous paragraph, we now consider the stochastic differential equation,

$$dx_t = f(x_t)dt + g(x_t)d\omega_t, \qquad \{x_t,\ t \geq 0,\ x\ \epsilon\ S_x \subset \mathbf{R}^n\}. \qquad (2.16)$$

where we assume $S_x$ is a open, simply connected subset of $\mathbf{R}^n$.

The question we ask is: are there diffusions of the type (2.16) which are equivalent ($C^\infty$ related) to a diffusion of the type

$$dz_t = Az_t dt + dw_t, \qquad \{z_t,\ t \geq 0,\ z\ \epsilon\ \mathbf{R}^n\}. \qquad (2.17)$$

Krener [23], Su [12] and Hunt, Su and Meyer [13] have answered related questions for deterministic nonlinear differential equations of the form $\dot{x} = f(x) + g(x)u(t)$. By properly interpreting the relation between (2.16) and the vector fields defined by the drift and diffusion terms we can extend their results to the s.d.e. case.  We proceed as follows:

We seek a transformation $\phi: S_x \to \mathbf{R}^n$ which is sufficiently smooth ($C^\infty$) so that $z_t = \phi(x_t)$ is a diffusion on $\mathbf{R}^n$ and satisfies Itô's rule.  That is

$$dz_t = \{\sum \frac{\partial\phi}{\partial x_i} f_i(x) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2\phi}{\partial x_j \partial x_i} g_i(x)g_j(x)\}\ dt +$$
$$\{\ \sum \frac{\partial\phi}{\partial x_i} g_i(x)\ dw_t\} \qquad (2.18)$$

and under which we have the equivalent linear stochastic differential equation

$$dz_t = Az_t dt + bdw_t \qquad (2.19)$$

for some (nxn) matrix A and (nxl) vector b.  We make the following assumptions: $\phi(0) = 0$, $\phi$ is 1-1 and invertible, $\phi$ is $C^\infty$ on $S_x$ and the pair (A,b) is controllable (in the usual sense of time invariant linear systems).

Now note the following result. If $\phi$ exists such that equations (2.18) and (2.19) are equivalent descriptions of the diffusion $\{z_t, t \geq 0\}$ then for $m = 1, 2, \ldots, n$

$$\sum_{j=1}^{n} g_j \sum_{i=1}^{n} \frac{\partial^2 \phi_m}{\partial x_j \partial x_i} g_i = -\sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial \phi_m}{\partial x_i} \frac{\partial g_i}{\partial x_j} g_j \qquad (2.20)$$

This is obtained by equating the diffusion terms in (2.18) and (2.19), differentiating with respect to $x_j$, multipling by $g_j$ and summing. If we interpret $g(x)$ as a vector field over $\mathbb{R}^n$, then in terms of coordinates

$$\underline{d}g(x) = \sum_{j=1}^{n} \sum_{i=1} \frac{\partial g_j}{\partial x_i} (dx^i \wedge \frac{\partial}{\partial x_j}) \;,$$

which is easily interpreted as a linear functional on $T(S_x)$. That is, dg acts on $T(S_x)$ by way of the pairing $\langle \cdot, \cdot \rangle$. For example,

$$\langle \underline{d}g, g \rangle =: (\frac{\partial g}{\partial x})g$$

$$\text{and } (\frac{\partial g}{\partial x})g = \sum_{i=1}^{n} (\sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} g_j) \frac{\partial}{\partial x_i} \quad \epsilon \; T(S_x) \qquad (2.21)$$

Using (2.20) and (2.21) and equating terms in (2.18) and (2.19), we arrive at the conditions for equivalence (using exterior differentiation, $\underline{d}$):
let $\phi = (\phi_1, \cdots, \phi_n)$ then

$$\langle \underline{d}\phi_m(x), Q(x) \rangle = \sum_{1}^{n} a_{mi} \phi_i(x)$$

$$\langle \underline{d}\phi_m(x), g(x) \rangle = b_m \qquad (2.22)$$

where $A = (a_{mi})_1^n$, $b = (b_1, \ldots, b_n)'$, prime denoting transpose, and $Q(x)$ signifies the vector field induced by

$$Q(x) = f(x) - \frac{1}{2} \frac{\partial g(x)}{\partial x} g(x).$$

Remarks:

(1) Equations (2.22) are the linearization conditions obtained for the
deterministic differential equation $\dot{x} = Q(x) + g(x) \dot{u}(t)$

(2) The functions $Q(x) = f(x) - \frac{1}{2} \frac{\partial g(x)}{\partial x} g(x)$ are precisely of the form
required to associate the s.d.e. (2.16) with its induced vector
fields on $T(S_x)$ [14]. As is well known, $Q(x)$ represents the vector
field induced by the Stratonovich form of Eqn. (2.16).

## Necessary and Sufficient Conditions for Local Equivalence

To solve the simultaneous partial differential equations (22) easily, we
need to apply the Frobenius Theorem [10]. To make use of this theorem we
need assume that the pair (A,b) is controllable.

Following the procedure of Section II, equations (2.22) in matrix form are

$$\left\langle \underline{d}\phi, \, Q \right\rangle = A\phi$$

$$(2.23)$$

$$\left\langle \underline{d}\phi, \, g \right\rangle = b$$

Defining the Lie derivative by $L_Q(g) = [Q,g]$ where $[\cdot,\cdot]$ is the Lie bracket,
and then iteratively

$$L_Q^m(g) = [Q, \, L_Q^{m-1}(g)] \qquad m = 0, 1, \ldots$$

$$L_Q^o(g) = g$$

and imitating the method of Su [12] we obtain the (necessary and sufficient)
equations for equivalence as,

$$\left\langle \underline{d}\phi, \, L(Q, \, g) \right\rangle = C(-A,b) \, .$$

$$(2.24)$$

Here $C(-A,b)$ is the controllability matrix of the pair $(-A,b)$, $L(Q,g)$ is the
matrix whose columns are given by the vector fields $(L_Q^o(g), \, \cdots, \, L_Q^{n-1}(g))$,
$\underline{d}\phi$ is the vector of 1-forms $(\underline{d}\phi_1, \, \underline{d}\phi_2, \, \cdots, \, \underline{d}\phi_n)'$ and we have generalized the
natural pairing, $\left\langle \cdot,\cdot \right\rangle$, to these "vectors" in the obvious way. In more explicit

notation, let $x = (x_1, \ldots, x_n)$ be a Euclidean coordinate system on $S_x$ and

$$\underline{dx} = (\underline{dx}_1, \underline{dx}_2, \cdots, \underline{dx}_n)^\prime$$

$$\frac{\partial}{\partial x} = (\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_n})$$

then $\left\langle \underline{dx}, \frac{\partial}{\partial x} \right\rangle = (\delta_{ij}) = I$, the nxn identity matrix and hence

$$(\frac{\partial \phi}{\partial x}) \quad dx, \frac{\partial}{\partial x} \quad L = C(-A, b)$$

$$(\frac{\partial \phi}{\partial x}) = \begin{bmatrix} \dfrac{\partial \phi_1}{\partial x_1} & \cdots \cdots & \dfrac{\partial \phi_1}{\partial x_n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \dfrac{\partial \phi_n}{\partial x_1} & \cdots \cdots & \dfrac{\partial \phi_n}{\partial x_n} \end{bmatrix} \qquad (2.25)$$

$$L = \begin{bmatrix} g_1, & [Q,g]_1 & \cdots \cdots & [Q,L_Q^{n-1}(g)]_1 \\ & \cdot & & \\ & \cdot & & \\ & \cdot & & \\ g_n, & [Q,g]_n & \cdots \cdots & [Q,L_Q^{n-1}(g)]_n \end{bmatrix}$$

The main integrability result of Differential Geometry, the Frobenius
Theorem, is noted for completeness:

THEOREM 2 Let D be a "differential system" of dimension p on an n-
dimensional differentiable manifold M. Let the vector fields $\{X_1, \cdots, X_n\}$
be linearly independent and let V(D) be the set of all vector fields, X,
such that $X(r) \in D(r) \, \forall \, r \in M$. Then D is completely integrable if and
only if V(D) is a Lie Algebra over F(M). Thus if $\{X_1, \cdots, X_p\}$ span

-22-

V(D) on a neighborhood U, then D is completely integrable if and only if there are ($C^\infty$) functions $C_{jk}^i : U \rightarrow R$ such that

$$[X_k, X_j] = \sum_1^p C_{jk}^i X_i, \quad 1 \le j, k \le p \quad (2.26)$$

In [12], this is applied to the deterministic linearization problem under feedback to show the existance of a diffeomorphism $\phi$ when,

(i)   $\{g, L_Q^1(g), \cdots, L_Q^{n-1}(g)\}$ are linearly independent in $T(S_x)$ in a neighborhood of the origin, U.

(ii)   $\text{Span}\{[L_Q^m(g), L_Q^i(g)], 0 \le m, j \le n-2\}$ is a Lie Algebra over $F(S_x)$ in $T(S_x)$ with $\{[g, \cdots, L_Q^{n-2}(g)\}$ in U.

(iii) $Q(o)=0$

In the present setting, a linear s.d.e. representation for (2.16) exists (locally) if the above conditions hold. Obviously, if these conditions hold on all of $T(S_x)$ then the equivalence is global.

In addition to the Frobenius Theorem, it is shown in [13] that a sufficient condition for $\phi$ to be a global transformation providing equivalence is,

(iv)   the Jacobian of $\phi$ satisfies the (noncharacteristic matrix) ratio conditions on $R^n$.

This assures that $\phi: S_x \rightarrow R^n$ is 1-1 on all of $R^n$.   The proof of the sufficiency of this condition is given in [13].

The necessary and sufficient condition (ii) is cumbersome to verify, when n is large.  A more practical test for this is available by using a version

of the Frobenius theorem stated in terms of the 1-forms dual to the set $L(Q,g)$. This is motivated in the discussion which follows.

Assume $L(Q,g)$ is a linearly independent set of vector fields on some $U \subset S_x$. Let $\omega_L = \{\omega_L^1, \cdots, \omega_L^n\}$ be the first dual to this set. Then

(a) $\omega_L$ is a vector whose components are 1-forms in $\Lambda^1(S_x)$

(b) Since a dual basis to $\{\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_1}\}$ on $U_*$ is $\{dx^1, \cdots, dx^n\}$, $\omega_L$ spans $U^* \subset \Lambda^1(S_x)$

(c) Using the matrix notation of (2.25), if

$$L(Q,g) = (L' \frac{\partial}{\partial x})' \qquad (2.27)$$

then we have the representation:

$$\omega_L = (L^{-1})\underline{dx}$$

(d) Since $L(Q,g)$ span $U_*$, $L^{-1}$ exists ($L$ is non-singular on all of $U$).

(e) Equation (2.24) is equivalent to the equation

$$\underline{d}\phi = C(-A, b)\omega_L \qquad (2.28)$$

Remarks:

(1) Equation (2.28) is a set of coupled partial differential equations for $\phi(x)$ (on $U$). Our goal is to find a solution of the form,

$$\phi(x) - \phi(0) = \int_{[o,x] \subset U} C(-A,b)\omega_L . \qquad (2.29)$$

This is a standard problem in the integration of differential forms on the manifold $\mathbf{R}^n$ (or $S_x$).

(2) Because $(A,b)$ is controllable, (2.19) can be chosen as a controller form. Then $C(-A,b)$ is upper triangular and $\phi_n(x)$ will depend only on $\omega_L^n$. Then remaining components, $\phi_{n-1}, \cdots, \phi_1$ can be built up from this solution. Hence the n equations can be solved one at a time, starting at n and working downwards, similar to the procedure in [12]. As will be seen in the sequel this is totally unnecessary as (2.29) can be integrated directly.

(3) If $\omega_L^n$ were exact then it is clear that $\phi$ would exist and equation (2.28) solved by solving the related equation $\omega_L^n = \underline{d}\theta$, where $\theta$ is a o-form. By Poincare's lemma, one would like to say that if $\underline{d}\omega_L^n = 0$ then $\exists\ \theta\ \varepsilon\ \Lambda^0(S_x)$ such that

$$\phi(x) = C(-A,b)\{\theta(x) - \theta(0)\} + \phi(0)$$

This is in fact true globally on $S_x$ (see the version of the converse to Poincare's lemma below). This leads us to the result stated in Theorem 4 below.

First we state a version of Frobenius Theorem in line with the kind of result we are looking for.

__THEOREM 3__ Let D be the differential system of dimension p on the n-dimensional manifold $S_x$ as before. Let $\omega_L = \{\omega_L^1,\ \cdots,\ \omega_L^n\}$ be the dual to $L(Q,g)$ on U then let $I(D) = \{$set of differential forms $\Omega|\Omega(r)\}$vanishes on $D(r)\ \forall\ r\ \varepsilon\ S_x\}$. Hence, on U, $\omega_L^{p+1},\ \cdots,\ \omega_L^n$ generate $I(D)$ and D is completely integrable if there exist $C^\infty$ functions $c_{jk}^i$ on $S_x$ such that

$$\underline{d}\omega_L^i = \sum_{p < j \le k} c_{jk}^i \omega_L^j \wedge \omega_L^k \qquad (2.30)$$

Proof: (see Sternberg [10]) On U*, since $L(Q,g)$ is a linearly independent set, we have uniquely the dual 1-forms (a basis on $\Lambda^1(S_x)$), $\{\omega_L^1, \cdots, \omega_L^n\}$ . On U*,$I(D)$ is generated by $\{\omega_L^{p+1},\ \cdots,\ \omega_L^n\}$ . As in Theorem 2, $\{X_1,\ \cdots,\ X_p\}$ span D. We have, using the relation from our preliminaries,

$$\left\langle X_i \wedge X_j | \underline{d}\omega_L^k \right\rangle = X_i \left\langle X_j | \omega_L^k \right\rangle - X_j \left\langle X_i | \omega_L^k \right\rangle - \left\langle [X_i, X_j] | \omega_L^k \right\rangle \qquad (2.31)$$

Taking i, j$\le$p, k>p, we have by applying definition (2.15),

$$\left\langle X_i \wedge X_j | d\omega_L^k \right\rangle = - \left\langle [X_i, X_j] | \omega_L^k \right\rangle \qquad (2.23)$$

and this vanishes if $[X_i, X_j]$ satisfies (2.26) hence the result.

Specializing this theorem to the problem at hand we have:

**THEOREM 4** The nondegenerate diffusion with $S_x = \mathbf{R}^n$ and completely specified by the s.d.e., $dx_t = f(x_t)dt + g(x_t)dw_t$, where f, g are $C^\infty$, have a controllable, linear s.d.e. representation (on $\mathbf{R}^n$) of the form,

$$dz_t = Az_t dt + bdw_t$$

if

(i)    $Q(0)=0$

(ii)   $L(Q,g)$ is a linearly independent set of vector fields on a neighborhood of the origin, $U \subset S_x$.

(iii)  the dual set of one forms to $L(Q,g)$ given by $\omega_L = \{\omega_L^1, \cdots, \omega_L^n\}$, has $d\omega_L^n = 0$ on all of U (i.e., $\omega_L^n$ is closed).


Proof:  Using the controllability assumption, pick a controller form for (A,b) then apply Theorem 3.

Remarks:

Global validity of the transformation, $\phi(x)$, is dealt with in [13] through study of the Jacobian of $\phi(x)$ and the ratio condition on the non-characteristic matrix; in essence, determining when $\phi$ is 1-1 on $\mathbf{R}^n$. A more succinct approach is argued as follows:

For $\phi(x)$ to exist, with arbitrary controllable pair (A,b), a necessary condition is that $\omega_L$ be a set of closed 1-forms (i.e., $\underline{d}\omega_L = 0$). The integrability question posed by (2.29) is then the same as the question: what conditions on (Q,g) assure that

$$\omega_L^n = \underline{d}\theta_n = \sum_1^n \frac{\partial \theta_n}{\partial x_j} dx^j$$

for some $\theta_n: \mathbf{R}^n \to \mathbf{R}$  (i.e., that $\omega_L^n$ is exact)?  When $L(Q,g)$ are linearly independent on the whole of $\mathbf{R}^n$, hence $\omega_L$ is a linearly independent set everywhere, then a global equivalence exists whenever $d\omega_L^n = 0$ by

virtue of the converse to Poincare's lemma (see Flanders [11], chapter 3). This gives the aforementioned simpler test in the following result.

__Corollary 1:__ Let the process $\{x_t, t \geq 0\}$, satisfying (2.16), have state space $S_x$ diffeomorphic to $\mathbf{R}^n$.

Suppose $L(Q,g)$ is a linearly independent set on all of $T(S_x)$ and let $\omega_L$ denote the dual to this set. Then a necessary and sufficient condition for the global equivalence of the s.d.e.s (2.16) and (2.17) is that $\underline{d}\omega_L^n = 0$ on all of $\Lambda^2(S_x)$ (i.e., for all $x \in S_x$ and any $X \in T(S_x)$).

__Proof:__ Because $S_x$ is diffeomorphic to $\mathbf{R}^n$, note that $S_x$ has a global chart, say $(\mu, U)$. Using the proof of Flanders for the lemma concerning the exactness of a p-form on a domain U, note that $\mathbf{R}^n$ can be deformed to a point $(0 \in \mathbf{R}^n)$ by $\tau(s, x^1, \cdots, x^n) = (sx^1, sx^2, \cdots, sx^n)$.

Using the induced mapping of forms, $\tau^*$, and $\omega_L^n = \sum_1^n \alpha_i(x^1, \cdots, x^n)dx^i$, we have

$$\tau^*(\omega_L^n) = \sum_1^n \alpha_i(sx^1, \cdots, sx^n)d(sx^i)$$

$$= \sum_1^n \alpha_i(sx^1, \cdots, sx^n)(sdx^i + x^ids), \quad s \in I, \ I = [0, 1]$$

Now define the mapping $K:\Lambda^1(I \times \mathbf{R}^n) \to \Lambda^0(\mathbf{R}^n)$ by

$$K(\alpha(s,x)dx^j) = 0$$

$$K(\alpha(s,x)ds) = \int_0^1 \alpha(s,x)ds$$

for all "smooth" functions $\alpha: I \times \mathbf{R}^n \to \mathbf{R}^n$.

So

$$K(\tau^* \omega_L^n) = \sum_{i=1}^n \int_0^1 x^i\alpha(sx^1, \cdots, sx^n)ds$$

Let $\theta_L^n = K(\tau \star \omega_L^n)$, then a simple calculation shows that since $d\omega_L^n = 0$,

$$(\frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i}) = 0$$

and

$$d\theta_L^n = \omega_L^n .$$

Remarks: (1) This corollary gives a simple constructive test for the existance of a linear s.d.e. representation, either globally or locally depending on the domain on which $L(Q,g)$ is non-singular. There does not appear to be any simple conversion of this condition to one directly upon $(Q,g)$. It is, however, straightforward once one verifies the linear independence of $L(Q,g)$. Furthermore, the global result may breakdown if $\{x_t, t \geq 0\}$ is a diffusion restricted to a manifold other than one globally diffeomorphic to $\mathbf{R}^n$ (e.g., the unit sphere $S^n \subset \mathbf{R}^n$). The problem here is that one really needs a global chart or coordinate system to make it work globally. The technical point being that the converse to Poincare's lemma fails except on such spaces; one then needs to consider the cohomology of 1-forms on $S_X$. Hence, the state space $S_X$ must be globally diffeomorphic to $\mathbf{R}^n$, possessing only <u>natural</u> boundaries. In particular, $S_X$ must be a simply connected open subset of $\mathbf{R}^n$. Otherwise, Theorem 3 can only be asserted in a open neighborhood of the origin; in geometric language on a single coordinate neighborhood of a chart. With the restriction that $S_X$ be diffeomorphic to $\mathbf{R}^n$ and has a single global chart then we need not specify in Theorem 4.

(2) If $S_x$ is only locally diffeomorphic to $\mathbf{R}^n$ (i.e., is a general differentiable manifold) then Theorem 4 is applied to each chart in the open covering of $S_x$ as a $C^\infty$-manifold. Thus, if the charts are labeled $(\mu_\alpha, U_\alpha)$, we must find a corresponding $\phi_\alpha = \phi \circ \mu_\alpha$ on $U_\alpha$. Piecing together these $\phi_\alpha$, if they exist, an equivalent linear s.d.e. can be found for each nonlinear s.d.e. representation of the diffusion on $S_x$. It may also turn out that a linear s.d.e. representation will exist only on a subset of all charts in the manifold covering.

## EXAMPLES

(1) Let $(x_t, y_t, t \geq 0)$ be a 2-dimensional process satisfying the s.d.e.

$$d\binom{x_t}{y_t} = \begin{bmatrix} \alpha x_t + \beta y_t e^{-y_t} - x_t^2 e^{y_t} \\ x_t e^{y_t} \end{bmatrix} dt + \begin{bmatrix} e^{-y_t} \\ 0 \end{bmatrix} dw_t .$$

Applying Theorem 4

$$Q(x,y) = \begin{bmatrix} \alpha x + \beta y e^{-y} - x^2 e^y \\ x e^y \end{bmatrix}$$

$$L(Q,g) = \begin{bmatrix} e^{-y} & e^{-y}(xe^y - \alpha) \\ 0 & -1 \end{bmatrix}$$

$$\det L(Q,g) = e^{-y}$$

$$\omega_L = \begin{bmatrix} e^y & (xe^y - \alpha) \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \omega_L^1 \\ \omega_L^2 \end{bmatrix}$$

and

$$\underline{d\omega}^2{}_L = 0$$

$$\underline{d\omega}^1{}_L = (e^y - e^y)dy \wedge dx = 0 .$$

So

$$\phi(x,y) = C(-A,b) \int \omega_L = C(-A,b) \begin{bmatrix} xe^y \\ -y \end{bmatrix}.$$

Then under the transformation $(u_t, v_t) = (x_t\, e^{y_t}, y_t)$ we have the equivalent linear representation.

$$d\begin{bmatrix} u_t \\ \dot{v}_t \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} dt + \begin{bmatrix} 1 \\ 0 \end{bmatrix} dw_t$$

Also, although clearly unlikely, if the accompanying observation were of the form

$$dz_t = (c_1\, x_t\, e^{y_t} + c_2\, y_t)dt + dv_t \quad,$$

then the conditional density for the process $(x_t, y_t)$ given $\{z_s,\ 0 \le s \le t\}$ could be determined from a Kalman filter for the linear representation of the problem.

(2)  An interesting example is obtained by generalizing example 3 of Section 2 to 2-dimensions.

Let $q_\pm(x,y) = (1 \pm x^2 \pm y^2)$   and set

$$f(x,y) = \begin{bmatrix} y\left(\dfrac{q_+(x,g) - 4x^2}{q_+(x,y)}\right) + \dfrac{q_-^2(x,y)}{q_+(x,y)}\left(4x^2\left(\dfrac{q_+(x,y) + 1}{q_+(x,y)} - 3\right)\right) \\[4mm] x\left(\dfrac{q_+(x,y) - 4y^2}{q_+(x,y)}\right) + \dfrac{y\, q_-^2(x,y)}{4q_+(x,y)}\left(q_-^2(x,y) - 4x^2\right) \end{bmatrix}$$

$$g(x,y) = \begin{bmatrix} \dfrac{q_+(x,y) - 2x^2}{q_+(x,y)}\, q_-(x,y) \\[5mm] -2xy\, \dfrac{q_-(x,y)}{q_+(x,y)} \end{bmatrix}$$

Then

$$d \begin{bmatrix} x_t \\ y_t \end{bmatrix} = f(x_t, y_t)dt + g(x_t, y_t)dw_t \quad , \quad \{w_t \in \mathbf{R}\} ,$$

with initial condition $x_o^2 + y_o^2 < 1$, with probability 1, defines a diffusion process on the unit disc, $D^2 = \{(x,y) | x^2 + y^2 < 1\}$ for all $t \geq 0$.

Since $D^2$ is diffeomorphic to $\mathbf{R}^2$ one would suspect that this process may have an equivalent linear s.d.e. description on $\mathbf{R}^2$. Indeed, after very tedious computations, Theorem 4 is verified such that $\underline{d\omega}_L(x,y) = 0$ $\forall (x,y) \in D^2$ and the desired transformation is found to be

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \frac{1}{q_-(x,y)}$$

analogous to example 3 of section 2. Under this transformation the linear s.d.e. on $\mathbf{R}^2$ is

$$d \begin{bmatrix} u_t \\ v_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} dt + \begin{bmatrix} 1 \\ 0 \end{bmatrix} dw_t$$

## 4. STANDARD GENERALIZATIONS

### Multi-Input Models

The counterpart to Theorem 4 for multiple independent Brownian motion driving terms follows directly from standard developments of multivariable linear system theory. The nonlinear s.d.e's to be dealt with have the form

$$dx_t = f(x_t)dt + \sum_{i=1}^{m} g^i(x_t)dw_{it} \; ; \qquad (2.33)$$

$\{w_{it}, \; t \geq 0\} \in R$, $i=1, \cdots, m$ ($m < n$) are mutually statistically independent, standard Brownian motions. $f(x)$, $g_1(x)$, ..., $g_m(x)$, are assumed $C^\infty$, Borel measurable mappings $(x) \in R^n \rightarrow R^n$ and the diffusion process $\{x_t, \; t \geq 0\} \in R^n$ satisfying (2.33) is assumed non-degenerate with only natural boundaries.

Again, as in the previous sections, the intent is to determine when a smooth mapping, T, exists which carries the nonlinear representation (2.33) into an equivalent linear, time-invariant s.d.e. representation:

$$dz_t = Az_t dt + Bdw_t \qquad (2.34)$$

where $w_t$ represents the (column) vector value Brownian motion with components $(w_{1t}, \ldots, w_{mt})$. $B = (b_1, \ldots, b_m)$ is an n x m matrix of constants representing the diffusion constants of the equivalent representation.

Because of the lack of a single unique choice for a linear realization in multi-input, multi-output systems, we need to choose a "preferred" form for the linear target pair (A,B). To make things concrete we choose the Popov form for (A,B), assuming of course that this realization is controllable. Also, the earlier remarks concerning the state space, $S_x$, of the nonlinear representation of the processes $\{x_t, \; t \geq 0\}$ carry over to this situation as well so there is no real restriction to assuming $S_x = R^n$.

where $v_{1t}$, $v_{2t}$ are statistically independent Brownian motions (standardized) as are $w_{1t}$, $w_{2t}$. The envelope, $r_t$, is a true envelope process in as much as,

$$r_t > 0 \; \forall_t \text{ with Prob. 1 if } r_0 > 0 \text{ with Prob. 1.}$$

The phase process, $\theta_t$, is a "3rd order" process: we model the "acceleration" in phase as white noise. This is well known to be a very reasonable model for uncertain phase in many applications without further apriori information. (See [1].) Indeed, under any reasonable physical interpretation the fluctuation in phase and frequency (i.e., uncertainty) is well modeled to first order as linearly dependent upon a small underlying random "phase acceleration".

With respect to the choice of model for the envelope, we remark that the bilinear form we propose allows a relatively wide variety of intuitively pleasing cases by varying the model parameters $\alpha$ and $\beta$ , which are viewed as fixed constants. For example a slowly fading or growing amplitude effect can be achieved (representing an advancing or receding source), as well as a constant power signal by virtue of the relation [3]:

$$E\{r^2(t)\} = E\{r_0^2\} \exp(2(\alpha - \tfrac{1}{2}\beta^2)t + 2\beta^2 E\{(v_{1t})^2\})$$

$$= E\{r_0^2\} \exp(2(\lambda + \beta^2)t)$$

This model also has another motivation which comes from the applicability of both the linear equivalence transformation theory and the IQ looked loop (IQLL) approach.

The appendix outlines the transformation theory and indicates its application to (3). Here we simply remark that the envelope, $r(t)$, has the representation

$$r_t = r_0 \, e^{(z_t + \lambda_t)} , \quad r_0 > o \text{ with probability 1}$$

where
$$\lambda = \alpha - \tfrac{1}{2}\beta^2 \text{ and } z_t = \int_0^t \beta \, dv_{1s}$$

By Ito's rule it is clear that $dr_t = \alpha r_t dt + \beta r_t dv_{1t}$

## 2. Signal & Noise Model

A variety of applications, such as communications, radar and sonar, deal extensively with a signal-in-noise representation of the form:

$$z(t) = r(t) \cos(\omega_0 t + \theta(t)) + n(t).$$

Here $r(t)$ is the envelope, $\omega_0$ a fixed (possibly unknown) carrier frequency, $\theta(t)$ a phase disturbance or uncertainty and $n(t)$ the (wideband) noise. Assuming $\omega_0$ is sufficiently known, after heterodyning and low pass filtering we typically have the "quadrature" observations:

$$y_1(t) = r(t) \cos \theta(t) + s \, w_1(t)$$
$$y_2(t) = r(t) \sin \theta(t) + s \, w_2(t). \tag{3.1}$$

We are assuming, of course, that $r$, $\theta$ are sufficently slowly time varying and that "s" represents the noise level after bandshifting.

In many signal processing situations, passive sonar in particular, this representation arises in an obvious way. Furthermore, in many ocean environments, $w_1(t)$ and $w_2(t)$ are modeled as statistically independent Gaussian noise processes to a very good approximation. This is the basic model we consider here.

For convenience we will henceforth deal with the "differential" form of the observations:

$$dy_t = d \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} r_t \cos \theta_t \\ r_t \sin \theta_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} \tag{3.2}$$

where $w_{1t}$, $w_{2t}$ are statistically independent Brownian motions. Accordingly, we propose as a state dynamical model the following:

$$dr_t = \alpha r_t \, dt + \beta r_t \, dv_{1t}$$

$$d \begin{bmatrix} \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} dt + \begin{bmatrix} 0 \\ 0 \\ q_2 \end{bmatrix} dv_{2t} \tag{3.3}$$

the original motivation for the transformation theory of the previous chapter. In particular, recent work [1,2] suggests that in a problem of signal phase and frequency tracking, EKF's formulated in terms of a new coordinate system for the state which causes the observation to become a linear functional on the state variables outperforms the usual EKF at low Signal-to-Noise ratio (SNR). We investigate the present conjecture in relation to these results.

In section 2 we define and justify dynamics and observation equations for our study. These are useful models for stochastic narrowband signals in additive noise. Throughout we have in the back of our mind the goal of detection of these narrowband sinusoid-like signals using the usual quasi-coherent approach. Section 3 discusses the discrete-time filtering algorithms we evaluate. The first algorithm is drawn from the work of Johnson, et. al. [1,2]; the so-called I-Q locked loop tracker. The second algorithm is based on the linear equivalence transformation of the previous chapter and the associated conjecture. In section 4 we present the results of a Monte Carlo analysis. The data presented suggests that, depending upon the SNR regime, the conjecture is indeed true, although not uniformly so. Finally, an Appendix outlines the transformation theory applied to the I-Q Locked Loop state equations studied in sections 2 and 3.

CHAPTER 3

## A Linear Equivalent Tracking Algorithm for the I-Q Locked Loop

### 1. Introduction

In tracking and recursive filtering problems, linear dynamics and linear (in the state) observations is the most desirable situation: the optimal solution is easily obtained by realizing a finite dimensional filter (Kalman-Bucy filter). The optimal solution behavior is well-understood, much analyzed, and computational procedures have been thoroughly documented. In the previous chapter it was shown how certain nonlinear dynamical models (of the state) can be transformed (pointwise) into wholly equivalent, completely controllable linear models. When the transformed observation system also happened to result in a set of equations linear in the new state variables, the optimal solution for this _linear_ system (of Kalman-Bucy type) could then be transformed back to the desired state estimates. This leads to a realization of a recursive state estimation solution for the original nonlinear stochastic differential equation (s.d.e.). Unfortunately, this rather special happenstance is rare. Thus the practical significance of the transformation theory comes into serious question. On the other hand, it appears to be relatively easy to find examples of practical signficance where the signal dynamics model indeed satisfy the conditions of the _linearizing_ transformation. Hence the transformation theory may have some practical applications.

In view of these observations, it is of some interest to ascertain if, in the case of these "closer to linear" problems, somewhat better behaved _approximate_ recursive filtering algorithms can be formulated. This Chapter investigates this conjecture: In the case where system dynamics admit an equivalent linear representation, irrespective of the nonlinearity in the corresponding observation, do Extended Kalman filters (EKF) based on the linear formulation evidence any improved performance over directly applied Extended Kalman Theory? Indeed this is

[14] R. W. Brockett, Nonlinear Systems and Differential Geometry, Proc. IEEE, Vol. 64, No. 1, 1976, pp. 61-72.

[15] R. Herman, Krenner, A., Nonlinear Systems and Differential Geometry, Proc. IEEE, Vol. 64, No. 1, 1976, pp. 61-72.

[16] Geometric Methods in Systems Theory, Mayne and Brockett, eds., Dorchecht, The Netherlands, D. Reidel

[17] R.W. Brockett, Nonlinear Systems and Nonlinear Estimation Theory, in Stochastic Systems: The Mathematics of Filtering and Identification and Applications, D. Reidel Publishing Co., 1981, pp. 441-477.

[18] R.W. Brockett, Clark, J.M.C., The Geometry of the Conditional Density Equations, Proc. Oxford Conf. on Stochastic Systems, Oxford, England, 1978.

[19] G.L Blankenship, C.H. Liu and S.I. Marcus, Asymptotic Expansions and Lie Algebras for some nonlinear filtering problems, IEEE Trans. Auto. Contr., AC-28, No. 7, 1983, pp. 787-797.

[20] H. Sussman, Approximate finite dimensional filters for some nonlinear problems, Stochastics, 7 (1982), pp. 183-203.

[21] W. Hopkins, Multiplicative asymptotic expansions for filtering linear diffusions with small polynomial nonlinearities, IEEE Conf. Decision and Contr. (1982), pp. 192-200.

[22] G.W. Blumen, On the transformation of diffusion processes into the Weiner process, SIAM J. Appl. Math., 39, No. 2, October 1980, pp. 238-247.

[23] A.J. Krener, On the equivalence of Control Systems and the linearization of nonlinear systems, SIAM. J. Control, 11, No. 4, November 1973, pp. 670-676.

Chapter 2 - References

[1] E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971

[2] V. Benes, Exact Finite Dimensional Filters for Certain Diffusions with Non-linear Drift, Stochastics; 1981, Vol. 5, pp. 65-92.

[3] D. Ocone, Nonlinear Filtering Problems with Finite Dimensional Estimation Algebras, Joint Auto. Cont. Conference, 1980, pp.

[4] R.S. Lipster and A.N. Shiryaev, Statistics of Random Processes II, Springer-Verlag, New York 1972

[5] W.M. Wohnam, Some Applications of Stochastic Differential Equation to Optimal Nonlinear Filtering, SIAM J., Ser. A, Vol. 2, No. 3, 1965, pp. 347-369.

[6] W.S. Wong, New Classes of Finite-Dimensional Nonlinear Filters, Systems and Cont., August 1983, pp. 155-164.

[7] R.W. Brockett, Classification and Equivalence in Estimation Theory, Proc. 18th IEEE Cont. Dec. and Control, Ft. Lauderdale, 1979, pp. 172-175.

[8] _____, The Global Description of Locally Linear Systems, in Springer-Verlag Lecture Notes on Control and Information Sciences, Vol. 29 (D. Hinrichsen and A. Isidori, eds.), 1982, pp. 1-8.

[9] J. Baras, Group Invariance Methods in Nonlinear Filtering of Diffusion Processes, Proc. 19th IEEE Conf. Dec. and Control, Albequerque, 1980, pp. 72-79.

[9a] V. Benes and I. Karatzas, Estimation and Control for Linear, partially observable systems with non-Guassian initial distribution, Stochastic Processes Appl. 14 (1983), pp. 233-248.

[9a] A. Makowski, Results on the filtering problem for Linear Systems with non-Gaussian Initial Conditions, Proc. 21st IEEE CDC, 1982, pp. 201-204.

[10] S. Sternberg, Lectures on Differential Geometry, Prentice-Hall, New York, 1962.

[11] H. Flanders, Differential Forms with Applications to the Physical Sciences, Academic Press, New York, 1963.

[12] R. Su, On the Linear Equivalent of Nonlinear Systems, Systems and Cont. Lett., Vol. 2, No. 1, July 1982, pp. 48-52.

[13] L.R. Hunt, Su, R., Meyer, G., Global Transformation of Nonlinear Systems, IEEE Trans. Auto. Cont., Vol. 28, No. 1, January 1983, pp. 24-30.

Combining this with (2.38) gives

$$\phi'(x)\{[Q,\sigma](x) + (f' \circ \phi(x))\sigma(x)\} = 0$$

$$\Longleftrightarrow [Q,\sigma](x) + (f' \circ \phi)(x)\sigma(x) = 0$$

Therefore, using (2.38) and substituting

$$\left(\frac{Q(x)}{\sigma(x)}\right)^2 + \frac{[Q,\sigma](x)}{\sigma(x)} = (a\phi(x))^2 + b\phi(x) + c - \phi^2(x)$$

Letting $V(x) = \left(\frac{Q(x)}{\sigma(x)}\right)^2 + \frac{[Q,\sigma](x)}{\sigma(x)}$

$$\phi(x) = \begin{cases} \dfrac{V(x) - c}{b} & , \quad a=1, \ b\neq 0 \\[3mm] \dfrac{-b\pm(b^2 - 4(a^2 - 1)(c - V(x))^{1/2}}{2(a^2 - 1)}, & (a\neq 1) \\[3mm] \displaystyle\int^x \frac{d\xi}{\sigma(\xi)} & a=1, \ b=0 \end{cases} \qquad (2.39)$$

Summarizing,

Equations (2.36) can be transformed to a Benes equation with drift $f(\cdot)$ iff

(1) $|\sigma(x)| > 0 \quad \forall x \in S_x$

(ii) for some at least twice continuously differentiable function $\phi(x)$

and scalars a,b & c,

$$\left(\frac{Q(x)^2}{\sigma(x)}\right) + \frac{[Q,\sigma](x)}{\sigma(x)} = (a^2 - 1)\phi^2(x) + b\phi(x) + c$$

In this case $\phi(x)$ is given by (2.39) and

$$f \circ \phi(x) = \left(\frac{Q(x)}{\sigma(x)}\right)$$

## Transformation of a Nonlinear s.d.e. to a Benes s.d.e.

Discussion is restricted to the scalar case. Consider the diffusion process satisfying the s.d.e.

$$dx_t = m(x_t)dt + \sigma(x_t)dw_t \qquad (2.36)$$

where $\{w_t, \; t \geq 0\}$ is a standard scalar Brownian Motion. Assume $m(x)$, $m'(x)$, $\sigma(x)$, $\sigma'(x)$ and $\sigma''(x)$ are continuous functions on $\mathbb{R}$ and $|\sigma(x)| > 0 \; \forall \; x \in \mathbb{R}$. The purpose here is to find a state variable transformation, $\phi$, such that $u_t = \phi(x_t)$ satisfies the Benes equation

$$du_t = f(u_t)dt + dw_t \qquad (2.37)$$

where $f(u)$ satisfies a Riccati differential equation

$$f'(u) + f^2(u) = (au)^2 + bu + c - u^2 \qquad (2.38)$$

a,b,c real constants.

Following the procedure used in Section 2 and equating like terms gives the equations

$$\phi'(x)Q(x) = f \circ \phi(x)$$

$$\phi'(x)g(x) = 1$$

Note the following identities

$$(\phi'(x)Q(x))' = f'(u)\phi'(x)\Big|_{u=\phi(x)}$$

$$(\phi'(x)\sigma(x))' = 0$$

$$(\phi'(x)\sigma(x))'Q(x) - (\phi'(x)Q(x))'\sigma(x)$$

$$= \phi'(x) \, [Q,\sigma](x) = - f'(u)\phi'(x)\Big|_{u=\phi(x)}\sigma(x)$$

$$= - f' \circ \phi(x)$$

**Theorem 5:** Let $L_k$ $(Q,g)$ represent the set of vector fields generated by the iterative bracketing operation in Section 3 on the vector fields associated with an n-dimensional s.d.e., as in (2.16). Assume $L_{n-1}(Q,g)$ does not satisfy the conditions of Theorem 4. Then there does not exist a mapping $\phi : \mathbb{R}^n \to \mathbb{R}^{n+m}$ $(m > 0)$ which linearizes the s.d.e..

**Proof:**

Since $L_{n-1}(Q,g)$ is singular, assume that for some $m > 0$, $L_{n+m-1}(Q,g)$ has full rank and let $\phi$ be an at least twice continuously differentiable map $(x) \in \mathbb{R}^n \to \mathbb{R}^{n+m}$. Assume that $z_t = \phi(x_t)$ satisfies a controllable linear s.d.e. in n+m dimensions with controllable realization (A,b) where

$$\left\langle d\phi, \ L_{n+m-1}(Q,g) \right\rangle = C(-A,b) \quad .$$

In matrix form

$$(\partial_x \phi) L_{n+m-1}(Q,g) = C(-A,b) \quad .$$

Since C(-A,b) is a controllability matrix of an n+m dimensional state space realization, rank C(-A,b) = n+m. Also, since $L$ is full rank

$$(\partial_x \phi) = C(-A,b) L^T (LL^T)^{-1}$$

Hence, $(\partial_x \phi)$ has rank at most n and therefore

$$\text{rank } C(-A,b) = \text{rank } (\partial_x \phi) L \leq n$$

$\Rightarrow$ the pair (-A,b) is not controllable and thus is a contradiction to the hypothesis.

## 5. RELATED MISCELLANEOUS TOPICS

The nonlinear stochastic differential equations which satisfy the foregoing transformation theory are a rather special subset of the problems one might encounter in practice. With a view to broadening the collection of s.d.e.'s which might be studied in the approximate filtering framework studied in the next chapter, this section considers two further questions:

(1) What are the comparable conditions to those of the previous sections with which to recognize when a given nonlinear s.d.e. is equivalent to some other, perhaps more useful, although nonlinear s.d.e.?

(2) Under conditions similar to those of the previous sections is it possible to find nonlinear s.d.e.'s which may linearized by a point-wise transformation to a higher dimension state space?

As to first question it is clear that, given a target s.d.e., the mapping from a general nonlinear s.d.e. to this form can be described in terms of a set of ordinary differential equations. Conditions for which this set of equations may be solved, if they are known, will describe the desired mapping. An example using the scalar Benes equation [2] is given in Appendix 2.1. The utility of this example is unknown.

With respect to the latter question, it is obvious that one can take an n-dimensional linear s.d.e. and map the corresponding state variables into another set of, say, n-m variables which satisfy a nonlinear s.d.e. Such a mapping is not 1-1 nor will it be invertible. In fact, the simple argument which follows shows that an n-dimensional nonlinear s.d.e. cannot be "controllably linearized" by any transformation to higher dimension if it does not already satisfy the conditions of Theorem 4.

$$
\omega_L(x) = \begin{bmatrix} \dfrac{x_1}{\beta(x_1^2 + x_2^2)} & \dfrac{x_2}{\beta(x_1^2 + x_2^2)} & 0 & 0 \\[4mm] 0 & 0 & 0 & 0 \\[4mm] 0 & 0 & -1 & 0 \\[4mm] \dfrac{-x_2}{\beta(x_1^2 + x_2^2)} & \dfrac{x_1}{\beta(x_1^2 + x_2^2)} & 0 & 0 \end{bmatrix}
$$

Since

$$
\frac{\partial}{\partial x_2} \left( \frac{x_1}{(x_1^2 + x_2)} \right) = \frac{\partial}{\partial x_1} \left( \frac{x_2}{(x_1^2 + x_2^2)} \right)
$$

it follows that $\underline{d}\omega_L = 0$ and therefore integration yields $\omega_L = \underline{d}\theta_L$ where

$$
\theta_L(x) = \text{col} \left( \frac{\ln(x_1^2 + x_2^2)}{2\beta}, \ x_4, \ -x_3, \ \text{Tan}^{-1}(\frac{x_2}{x_1}) \right).
$$

A version of the desired transformation is then

$$
T(x) = \text{col} \left( \frac{\ln(x_1^2 + x_2^2)}{2\beta}, \ \text{Tan}^{-1}(\frac{x_2}{x_1}), \ x_3, \ x_4 \right)
$$

with the corresponding equivalent linear s.d.e. representation

$$
dz_t = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} z_t dt + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix}
$$

This example is used in Chapter 3 to study an approximate filtering application of the transformation theory developed in this chapter.

where $\beta \neq 0$. Applying corollary 2 we have

$$Q(x) = \begin{bmatrix} \lambda x_1 - x_3 x_2 \\ \lambda x_2 - x_3 x_1 \\ x_4 \\ 0 \end{bmatrix} \qquad g_1(x) = \begin{bmatrix} \beta x_1 \\ \beta x_2 \\ 0 \\ 0 \end{bmatrix} \qquad g_2(x) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where $\alpha - \frac{1}{2}\beta^2$.

$$[Q, g_1](x) = 0 \quad (\Rightarrow L_Q^m(g_1) = 0 \;\; \forall \;\; m > 1)$$

$$[Q, g_2](x) = \text{col } (0,0,-1,0)$$

$$L_Q^2(g_2)(x) = \text{col } (-x_2, x_1, 0, 0)$$

$$L_Q^3(g_2)(x) = 0$$

Hence, the Kronecker indices are $k_1 = 1$, $k_2 = 3$ and a representation for the vector fields $L_2$ is,

$$L_2(x) = \begin{bmatrix} \beta x_1 & 0 & 0 & -x_2 \\ \beta x_2 & 0 & 0 & x_1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad .$$

$L_2(x)$ is full rank everywhere except at $x_1 = x_2 = 0$. If the diffusion

process represented by (2.35) is assumed to be generated from an orthogonal

representation of a "narrowband" signal, with $x_{1t}$ and $x_{2t}$ representing the

quadrature components (c.f. Chapter 3, section 3), then it follows that

$x_{1t}^2 + x_{2t}^2 > 0$ for all $t \geq 0$. With this $L_2(x)$ is everywhere ¹ rank

(on the state space) and the coefficients of the first dual to $L_2(\cdot)$ are, in

matrix form,

forms $\underline{d}T$ with the vector fields $L_Q^{\ell_j}(g_j)$, $j=1, \cdots, m$, $\ell_j = 0$, $1, \cdots, k_j$ and use of a Popov form for the pair $(A,B)$ with Kronecker indices $(k_1, \cdots, k_m)$.

(2) The strong conditions in (ii), $k_j > 0$ and at least one $j$ for which $k_j > 1$, can be relaxed provided that for all $j$ such that $k_j = 0$, $g_j(x)$ may be represented by

$$g_j(x) = \sum_{\{\ell \mid k_\ell \neq 0\}} \alpha_\ell g_\ell(x), \qquad \alpha_\ell \text{ } \underline{\text{scalars}}.$$

### An Example

The following example is motivated by a realistic signal tracking problem studied in the next chapter. The nonlinear state equations are as follows:

Let $x_t = \text{col}\,(x_{1t}, x_{2t}, x_{3t}, x_{4t}) \in \mathbb{R}^4$,

$$dx_t = \begin{bmatrix} \alpha x_{1t} - x_{3t} & x_{2t} \\ \alpha x_{2t} - x_{3t} & x_{1t} \\ x_{4t} \\ 0 \end{bmatrix} dt + \begin{bmatrix} \beta x_{1t} & 0 \\ \beta x_{2t} & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} \qquad (2.35)$$

With these preliminaries the multi-input result is stated in the following
Corollary to Theorem 4:

$\underline{\text{Corollary 2}}$: Let $\{x_t, \ t \geq 0, \ x \in \mathbf{R}^n\}$ be the non-degenerate diffusion
process satisfying the non-linear s.d.e. (2.33). Define

$$Q(\underline{x}) = f(\underline{x}) - \frac{1}{2} \sum_1^m \left\langle \underline{d}g_i(\underline{x}), \ g_i(\underline{x}) \right\rangle \ .$$

There exists a simply connected open domain, $U \subset \mathbf{R}^n$, and a smooth
transformation $T: U \to \mathbf{R}^n$ such that $z_t = T(x_t)$ has the representation (2.34)
if and only if:

(i)  $Q(o) = 0, \ 0 \in U$

(ii)  there exist positive integers, $k_j > 0$, $j=1, \ 2, \ \ldots, \ m$,

such that,

$$\sum_1^m k_j = n$$

with at least one $1 \leq j \leq m$  such that $k_j > 1$ and

$$L_m = \{g_1, \ \cdots, \ L_Q^{k_1-1}(g_1), \ \cdots, \ g_m, \ \cdots, \ L_Q^{k_m-1}(g_m)\}$$

form a linearly independent set on a simply connected open domain
$U \subset \mathbf{R}^n$ containing the origin

(iii) the duals to $L_m$, $\omega_L = (L_m^{-1})\underline{dx}$, are closed on U (i.e., $\underline{d}\omega_L = 0 \ \forall \in U$).

The required transformation, T, in this case is obtained by choosing
a Popov form (A,B) with Kronecker indices $k_j$, $j=1, \ 2, \ \ldots, \ m$,
computing the controllability matrix for $(-A,B)$ and performing the
integration

$$T(x) = \int_{(0,x) \subset U} C(-A,B)\omega_L \ .$$

Remarks:

(1) The proof of Corollary 2 follows directly from Theorem 4
after identification of the required inner products of the 1-

Under the IQLL framework the coordinate representation of interest is that which causes the observation to depend linearly upon the "signal". Therefore,

$$i_t = r_t \cos \theta_t$$

$$q_t = r_t \sin \theta_t$$

whence,

$$dy_t = \begin{bmatrix} i_t \\ q_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} . \qquad (3.4)$$

The state dynamics equations therefore become nonlinear in the new state variables:

$$d \begin{bmatrix} i_t \\ q_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} \alpha i_t - \omega_t q_t \\ \alpha q_t + \omega_t i_t \\ \omega_t \\ 0 \end{bmatrix} dt + \begin{bmatrix} \beta i_t & 0 \\ \beta q_t & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix} \qquad (3.5)$$

Note that even when $\alpha = 0$, the case studied in [1], (3.5) would remain in coupled bilinear form.

From the appendix, the equivalent form for the "log normal" diffusion model given by (3.5) is

$$d \begin{bmatrix} z_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} dt + \begin{bmatrix} \beta & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix} \qquad (3.6)$$

which is completely linear in the (new) state coordinates $(z, \theta, \omega, \xi)$. Of course here the observation remains nonlinear in the states and is seen to be

$$dy_t = \begin{bmatrix} r_o\, e^{(z_t + \lambda t)} \cos \theta_t \\ r_o\, e^{(z_t + \lambda_t)} \sin \theta_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} \tag{3.7}$$

These two alternative representations provide the starting point for our comparison of approximate filtering algorithms in response to the conjecture posed in section 1. The next section develops appropriate discrete-time approximations and discusses the Monte Carlo analysis set-up.

## 3. Development of the Filtering Algorithms

We restate the partial observation/filtering problem:

State Dynamics:

$$d \begin{bmatrix} r \\ \theta \\ \omega \\ \xi \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} dt + \begin{bmatrix} \beta r_t & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix} \qquad (3.8)$$

Observation:

$$d \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} r_t \cos \theta_t \\ r_t \sin \theta_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} \qquad (3.9)$$

where $(r_0, \theta_0, \omega_0, \xi_0)$ are statistically independent of $w_{1t}$, $w_{2t}$, $v_{1t}$, $v_{2t}$, $t \geq 0$ and the latter are statistically independent Brownian motions (standardized). We assume that $r_t$, $\theta_t$, $\omega_t$ and $\xi_t$ are statistically independent of the observation noises, $t \geq 0$.

Under the two alternative approaches, discussed in section 2., one can now formulate two extended Kalman filters. From the linear equivalent system the nonlinearity of the problem is isolated in the observation (c.f., eqns. 3.6 & 3.7) whereas in the IQLL coordinates 1 the nonlinearity has been isolated in the state dynamics equations (c.f., eqns (3.4) & (3.5)). The corresponding extended Kalman filters for these two representations are as follows:

(a)  Linear Equivalent Form

$$d \begin{bmatrix} \hat{z}_t \\ \hat{\theta}_t \\ \hat{\omega}_t \\ \hat{\xi}_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{z}_t \\ \hat{\theta}_t \\ \hat{\omega}_t \\ \hat{\xi}_t \end{bmatrix} dt + K_t^* \begin{bmatrix} dy^1 - r_o e^{\hat{z} + \lambda t} \cos \hat{\theta}\, dt \\ dy^2 - r_o e^{\hat{z} + \lambda t} \sin \hat{\theta}\, dt \end{bmatrix}$$

or more compactly,

$$d\hat{x}_{1e} = A\hat{x}_{1e}dt + K_t^* (dy - h(\hat{z}, \hat{\theta})\, dt)$$

(b) IQLL Form

$$d \begin{bmatrix} \hat{I}_t \\ \hat{Q}_t \\ \hat{\omega}_t \\ \hat{\xi}_t \end{bmatrix} = \begin{bmatrix} -\hat{\omega}_t \hat{Q}_t \\ \hat{\omega}_t \hat{I}_t \\ \hat{\xi}_t \\ 0 \end{bmatrix} dt + K_{IQ(t)} \begin{bmatrix} dy_t^1 - \hat{I}_t dt \\ \\ dy_t^2 - \hat{Q}_t dt \end{bmatrix}$$

The gain matrix $K_t^*$ is taken from the steady state solution of the Riccati equation for the estimated covariance of the linearized problem. Specifically,

$$\dot{P}_t = A P_t + P_t A^T + BB^T - P_t H^T(\hat{z}, \hat{\theta}) R^{-1} H(\hat{z}, \hat{\theta}) P_t$$

where $H(\hat{z}, \hat{\theta}) = H(\hat{x}_{1e}) \triangleq \begin{bmatrix} \hat{r} \cos \hat{\theta} & -\hat{r} \sin \hat{\theta} & 0 & 0 \\ \hat{r} \sin \hat{\theta} & \hat{r} \cos \hat{\theta} & 0 & 0 \end{bmatrix}$

$$\hat{r}_t \triangleq e^{\hat{z}_t + \lambda t}$$

and

$$BB^T \triangleq \begin{bmatrix} \beta^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & q_2^2 \end{bmatrix}$$

and $R^{-1} = \dfrac{1}{s^2}$  I (I is the 2x2 identify matrix)

It is easy to see that this equation decouples into a 1x1 equation for the "envelope" variance and a 3x3 equation for the "phase error" covariances. The steady state solutions are given by P* (where * denotes steady state):

$$P_z^* \overset{\Delta}{=} P_r^r = P_n^* = \frac{|\cdot \beta s|}{r}$$

$$P_\theta^* = \frac{s^2}{\hat{r}^2} \begin{bmatrix} 2\alpha & 2\alpha^2 & \alpha^3 \\ 2\alpha^2 & 3\alpha^3 & 2\alpha^4 \\ \alpha^3 & 2\alpha^4 & 2\alpha^5 \end{bmatrix} \quad \text{where } \alpha \overset{\Delta}{=} \left[ \frac{q_2^2 \, \hat{r}^2}{s^2} \right]^{1/6}$$

Finally,

$$K^* = P^* H^T (X_{LE}) R^{-1} = \begin{array}{l} \dfrac{P_{11}^*}{s^2} \quad \hat{r}(\cos\hat{\theta},\ \sin\hat{\theta}) \\[2mm] \dfrac{2\alpha}{s^2 S^2} \quad \hat{r}(-\sin\hat{\theta},\ \cos\hat{\theta}) \\[2mm] \dfrac{2\alpha^2}{s^2 S^2} \quad \hat{r}(-\sin\hat{\theta},\ \cos\hat{\theta}) \\[2mm] \dfrac{\alpha^3}{s^2 S^2} \quad \hat{r}(-\sin\hat{\theta},\ \cos\hat{\theta}) \end{array}$$

and $\quad K_{IQ} = J(\hat{X}_{IQ}/\hat{X}_{LE}) K^*$

where $J(\hat{X}_{IQ}/\hat{X}_{LE})$ represents the Jacobian matrix of the transformation from $\hat{X}_{LE}$ coordinates to the IQ coordinates and is given by

$$J(\hat{X}_{IQ}/\hat{X}_{LE}) = \begin{bmatrix} \hat{I} & -\hat{Q} & 0 & 0 \\ \hat{Q} & \hat{I} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

for $\hat{I}_t \cong e^{\hat{z}t + \lambda t} \cos\hat{\theta}_t$, $\hat{\theta}_t \approx e^{\hat{z}t + \lambda t} \sin\hat{\theta}_t$. From this

-51-

$$
K_{IQ} = \begin{bmatrix} \dfrac{2\alpha}{s^2 S^2}\ \hat{r}^2 + \gamma\hat{I}^2 & \gamma\hat{I}\hat{Q} \\[2ex] \gamma\hat{I}\hat{Q} & \dfrac{2\alpha}{s^2 S^2}\ \hat{r}^2 + \gamma\hat{Q}^2 \\[2ex] \dfrac{2\alpha^2}{s^2 S^2}\ \hat{Q} & \dfrac{2\alpha^2}{s^2 S^2}\ \hat{I} \\[2ex] \dfrac{\alpha^3}{s^2 S^2}\ \hat{Q} & \dfrac{\alpha^3}{s^2 S^2}\ \hat{I} \end{bmatrix}
$$

Above, the expressions used to simplify notation are

$$
S^2 = \hat{r}^2 \big/ s^2 \quad \text{and} \quad \gamma = \frac{1}{s^2}\left\{ P^*_{11} - \frac{2\alpha}{S^2} \right\}.
$$

In the IQLL filter in [1], the apparent "small" signal approximation is introduced to decouple the IQLL estimation equations as follows:

$$
P^*_{11} \cong 2\alpha / s^2 \qquad \text{for } \alpha << 1, \ (\hat{r}^2 << \frac{s^2}{q^2}) .
$$

Using this, and the gain matrices above, the two algorithms are,

Linear Equivalent:

$$
d\hat{z}_t = \frac{|\beta|}{s}\left\{ \cos\hat{\theta}_t\ dy_{1t} + \sin\hat{\theta}\ dy_{2t} - \hat{r}_t\ dt \right\}
$$

$$
d\begin{bmatrix} \hat{\theta}_t \\ \hat{\omega}_t \\ \hat{\xi}_t \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\theta}_t \\ \hat{\omega}_t \\ \hat{\xi}_t \end{bmatrix} dt + \begin{bmatrix} 2\alpha \\ 2\alpha^2 \\ \alpha^3 \end{bmatrix} \frac{\{-\sin\hat{\theta}_t\ dy_{1t} + \cos\hat{\theta}_t\ dy_{2t}\}}{\hat{r}_t}
$$

$$
\hat{r}_t \cong e^{\hat{z}_t + \lambda t}
$$

<div align="right">(3.10)</div>

IQLL:

$$d\begin{bmatrix} \hat{I}_t \\ \hat{Q}_t \end{bmatrix} = \begin{bmatrix} -\hat{\omega}_t \hat{Q}_t \\ \hat{\omega}_t \hat{I}_t \end{bmatrix} dt + 2\alpha \begin{bmatrix} dy_{1t} - \hat{I}_t dt \\ dy_{2t} - \hat{Q}_t dt \end{bmatrix}$$

$$d\begin{bmatrix} \hat{\omega}t \\ \hat{\xi}t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\omega}t \\ \hat{\xi}t \end{bmatrix} dt + \begin{bmatrix} 2\alpha^2 \\ \alpha^3 \end{bmatrix} \frac{\{-\hat{Q}_t dy_{1t} + \hat{I}_t dy_{2t}\}}{\hat{r}_t^2}$$

$$\hat{r}_t^2 = \hat{I}_t^2 + \hat{Q}_t^2 \tag{3.11}$$

## 4. Filter Comparisons

The two filtering algorithms were compared using Monte Carlo simulations. Each filter was given identical input data generated according to the discrete time approximation to (3.8) and (3.9). Simulations were performed for of signal-to-noise ratios ranging from -13B to +20dB. For the purpose of comparison, signal to noise ratio was computed as the integrated average signal power over the sample contained in the observation (3.9) divided by the noise variance

$$SNR = \frac{E\{\frac{1}{T}\int_{t_1}^{t_1+T} r^2(t)dt\}}{2s^2}$$

In Eqs (3.8) and (3.9), signal dynamics and observations were approximated using a 4th order stochastic Runge-Kutta scheme [6] at a sample spacing of 0.05 seconds.

Because the bilinear Ito equation,

$$dr = \alpha r dt + \beta r dv_{1t}$$

can be solved explicitly $(r_t = r_o e^{\beta v_{1t} + \lambda t}, \lambda = \alpha - \frac{1}{2}\beta^2)$

this solution was used to generate the sample paths for $r_t$ ($r_n = r_{tn}$, n = 1,2,...). The phase variables employed the stochastic integration routine.

Monte Carlo runs consisted of generating sample paths of length 240 seconds each using a white Gaussian random number generator whose correlation time corresponded to $2^{31}$ samples (i.e., the pseudo-random numbers are guaranteed not to repeat identically over any contiguous set of $2^{31}$ values). A corresponding ensemble of sample paths was produced for each signal in noise (SNR, etc.) condition. Each such ensemble was then processed using a discrete-time approximation to the two filtering algorithms (a predictor-corrector form in accordance with [1]). The output from the filtering routines were estimates of the original states $(\hat{r}_n, \hat{\theta}_n, \hat{\omega}_n, \hat{\xi}_n)$; n = 1,2,... . The filter implementations are given as follows:

Linear Equivalent Form:

$$\tilde{Z}n = \hat{Z}_{n-1}$$

$$\tilde{\theta}n = \hat{\theta}_{n-1} + \Delta \cdot \hat{\omega}_{n-1} + \frac{\Delta^2}{2} \hat{\xi}_{n-1}$$

$$\tilde{\omega}n = \hat{\omega}_{n-1} + \Delta \cdot \hat{\xi}_{n-1} \left. \right\} \quad \text{prediction}$$

$$\tilde{\xi}n = \hat{\xi}_{n-1} \ \& \ \tilde{r}_n = \hat{r}_{n-1} \, e^{\lambda\Delta}$$

$$\hat{Z}n = \tilde{Z}_n + \frac{|\beta|}{s} [\cos\tilde{\theta}_n \, \delta y_n^1 + \sin\tilde{\theta}_n \, \delta y_n^2 - \tilde{r}_n \Delta]$$

$$\hat{\theta}n = \tilde{\theta}_n + 2\tilde{\alpha}_n [-\sin\tilde{\theta}_n \, \delta y_n^1 + \cos\tilde{\theta}_n \, \delta y_n^2]/\tilde{r}_n$$

$$\hat{\omega}n = \tilde{\omega}_n + 2\tilde{\alpha}_n^2 [-\sin\tilde{\theta}_n \, \delta y_n^1 + \cos\tilde{\theta}_n \, \delta y_n^2]/\tilde{r}_n \left. \right\} \quad \text{correction}$$

$$\hat{\xi}n = \tilde{\xi}_n + \frac{(q_2)}{s} \frac{[-\sin\tilde{\theta}_n \, \delta y_n^1 + \cos\tilde{\theta}_n \, \delta y_n^2]}{\tilde{r}_n} \ \& \ \hat{r}_n = \tilde{r}_n \, e^{\hat{Z}_n}$$

where

$$\delta y_n^1 = y_n^1 - y_{n-1}^1, \ \delta y_n^2 = y_n^2 - y_{n-1}^2$$

and

$$\tilde{\alpha}_n = \left(\frac{q_2 \tilde{r}_n}{s}\right)^{1/3}$$

IQLL Form:

$$\begin{bmatrix} \tilde{I}n \\ \tilde{Q}n \end{bmatrix} = \begin{bmatrix} \cos\overline{\Delta\theta} & -\sin\overline{\Delta\theta} \\ \sin\overline{\Delta\theta} & \cos\overline{\Delta\theta} \end{bmatrix} \begin{bmatrix} \hat{I}_{n-1} \\ \hat{Q}_{n-1} \end{bmatrix} \left. \right\}$$

$$\begin{bmatrix} \tilde{\omega}_n \\ \tilde{\xi}_n \end{bmatrix} = \begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\omega}_{n-1} \\ \hat{\xi}_{n-1} \end{bmatrix} \qquad \text{prediction}$$

$$\begin{bmatrix} \hat{I}n \\ \hat{Q}n \end{bmatrix} = \begin{bmatrix} \tilde{I}n \\ \tilde{Q}n \end{bmatrix} + 2\tilde{\alpha}_n \begin{bmatrix} \delta y_n^1 - \tilde{I}_n \Delta \\ \delta y_n^2 - \tilde{Q}_n \Delta \end{bmatrix} \left. \right\}$$

$$\begin{bmatrix} \hat{\omega}_n \\ \hat{\xi}_n \end{bmatrix} = \begin{bmatrix} \tilde{\omega}_n \\ \tilde{\xi}_n \end{bmatrix} + \begin{bmatrix} 2\tilde{\alpha}_n^2 \\ \tilde{\alpha}_n \end{bmatrix} \{-\tilde{Q}_n \, \delta y_n^1 + \tilde{I}_n \, \delta y_n^2\}/\tilde{r}_n^2 \qquad \text{correction}$$

$$\tilde{\alpha}_n = \left(\frac{q_2^2 \, \tilde{r}_n^2}{s^2}\right)^{1/6}, \ \tilde{r}_n^2 = \tilde{I}_n^2 + \tilde{Q}_n^2$$

$$\hat{r}_n = (\hat{I}_n^2 + \hat{Q}_n^2)^{1/2}$$

-55-

and $\overline{\Delta\theta_n} = \tilde{\theta}_n - \hat{\theta}_{n-1} = \Delta \cdot \hat{\omega}_{n-1} + \frac{\Delta^2}{2} \hat{\xi}_{n-1}$ .

For the IQLL filter, to compare phase estimates between the two algorithms, $\hat{\theta}_n$ was taken as

$$\hat{\theta}_n = \tilde{\theta}_n + \overline{\Delta\theta}_n + 2\tilde{\alpha}_n \left[ -\tilde{Q}_n \, \delta y_n + \tilde{I}_n \delta y_n^2 \right] / \, \tilde{r}_n^2 \quad ,$$

since the IQLL does not produce these directly.

## Mean Square Error Comparison

The Monte Carlo comparison of the two filtering algorithms are summarized below. The comparisons are based on simulations for which the envelope power, $E\{r_t^2\}$, was time independent; for this situation choose $\alpha, \beta$ such that $\lambda = -\beta^2$. This allowed an essentially stationary representation of signal-to-noise ratio thoughout the numerical comparisons.

Figures 3-1 thru 3-4 show typical results for the two algorithms for estimation error (mean square) as a function of input SNR. Only the results for envelope and phase are shown as both algorithms perform equivalently in terms of mean square error for frequency ($\omega$) and frequency rate ($\xi$). The IQLL algorithm did show up to 0.5 dB improvement in frequency estimation error in a few cases, but not consistently.

Figures 3-1 and 3-2 show the typical situation for the phase and envelope estimates respectively. These figures also show the amplitude known (Figure 3-1) and phase known* (Figure 3-2) performance, respectively, for "lower bound" comparisons. These are the best possible estimates to be achieved from either algorithm in these situations. In both instances, the linear equivalent approximation appears, generally, to outperform the IQLL filter and performance advantage

Note: In Figure 3-2, the phase known case represents the optimal envelope estimation performance; the linear equivalent approach reduces to the optimal estimate for $r_t$ in this situation (within experimental and discretization error), whereas the IQLL formulation does not.

is clearly a function of input SNR. The best advantage for phase, within the region of acceptable error, ($<$ 1 rad), appears to be less than about 5-7dB and consistently occurs just above the transition SNR. The error advantage of the linear equivalent algorithm for envelope estimates can be quite large being a function of SNR and envelope state noise power. Figure 3-2 shows up to about 3dB improvement in estimation error which appears to be the typical improvement over the range of reasonable dynamics noise levels.

Figure 3-5 presents an aplications oriented comparison of the phase estimation performance of the two algorithms. This figure plots the input SNR at which the phase errors are below an acceptance threshold of 30° RMS. Although 30° is arbitrary, it is a reasonable criterion; errors such in excess of this value make the phase estimates essentially useless for any tracking/detection scenario. The independent variable in the figure is taken to be the phase dynamics noise level, $q_2$ in egn. (3.8); $|q_2|$ = 0 represents the phase known case and increasing $|q_2|$ represents, in a loose sense, the degree of inherent randomness in the phase process. With this interpretation, figure 3.5 indicates about a 1dB "threshold" advantage for the linear equivalent algorithm over the range of $|q_2|$ studied.

Finally, we note that at and below transition, results tend to be mixed, perhaps reflecting large statistical uncertainty associated with low SNR and large errors. Additionally, as SNR decreases, the linear equivalent algorithm diverges more rapidly than the IQLL. This reflects, perhaps, dominating effects of the observation nonlinearity of the linear equivalent algorithm in this regime of error (and SNR).

## Sensitivity to Initial Conditions

Experiments performed to discover the sensitivity of the two algorithms to the assumed initial state are summarized in Table 3-1. This data, in Table 3-1 (a) – (d), indicates the approximate time for convergence of the filter state estimates under the conditions of randomized initial state estimates. The controlling parameter in these data is the initial state estimate variance for each component, $(r_o, \theta_o, \omega_o, \xi_o)$, relative to the actual value. When convergence was not achieved after 4 minutes the filter was assumed not to converge. Note that although one state estimate does not converge the others may.

The IQLL algorithm was much less sensitive to initial amplitude than the Linear Equivalent; whereas exactly the opposite was the case for initial phase estimates. Frequency and frequency rate sensitivities appear to be approximately equivalent for the two filters. The sensitivity of the IQLL to initial phase estimation accuracy would seem to be the consequence of the use of I,Q instead of a phase coordinate directly and the fact that, for both algorithms, initial amplitude error appears to be relatively insignificant (e.g., $\sigma r_o^2 \approx 160.0$ and 40.0). This is also probably related to the "observability" of the $\theta$ coordinate. The robustness of the Linear Equivalent algorithm to initial phase estimate appears to be a consequence of the coordinates using the phase variable directly in estimation.

As a final note, these sensitivity analyses were performed such that individual state variable effects were "isolated". That is when one state initial condition was varied, the other state variable's initial conditions were fixed at the a priori known values. Composite effects are less clear but seem to be characterized by an ellisoidal region defined by the "isolated" variable sensitivity boundaries. (c.f. Table 3-2).
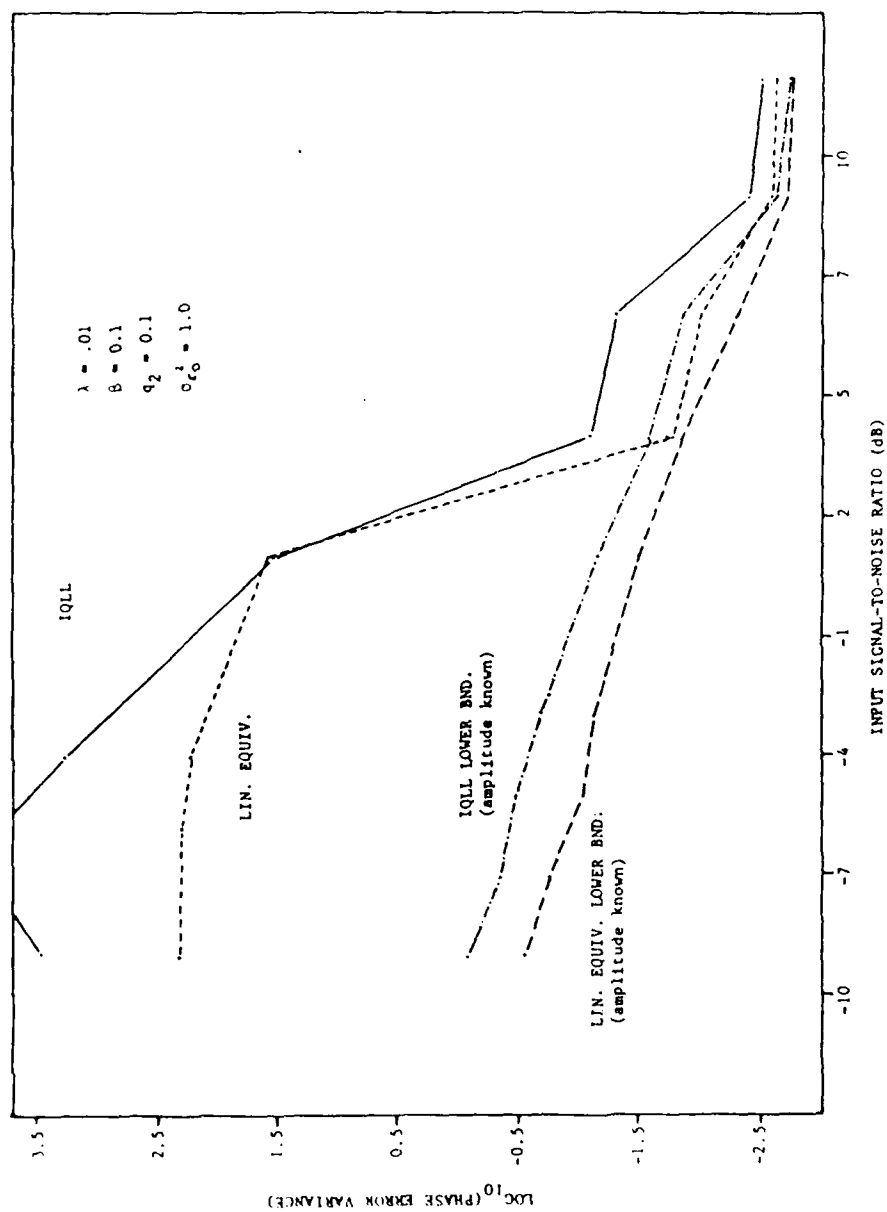
Figure 3-1 MS Phase Estimation Error Comparison; Linear Equivalent versus IQLL Approximate Filters with amplitude known lower bounds ($\lambda = -0.01$, $\beta = 0.1$, $q_2 = 0.1$)
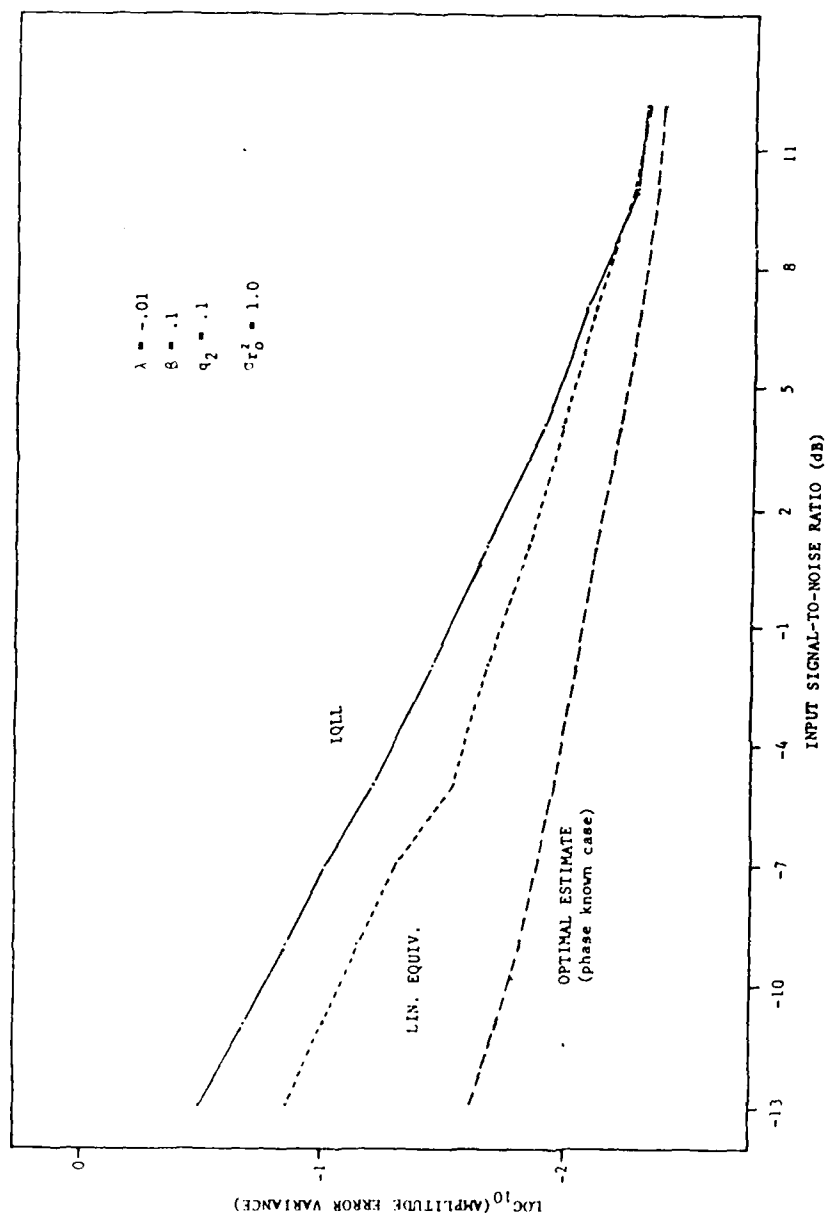
Figure 3-2 MS Amplitude Estimation Error Comparison; versus IQLL Approximate Filters with phase known lower bounds ($\lambda = -.01$, $\beta = 0.1$, $q_2 = 0.1$)
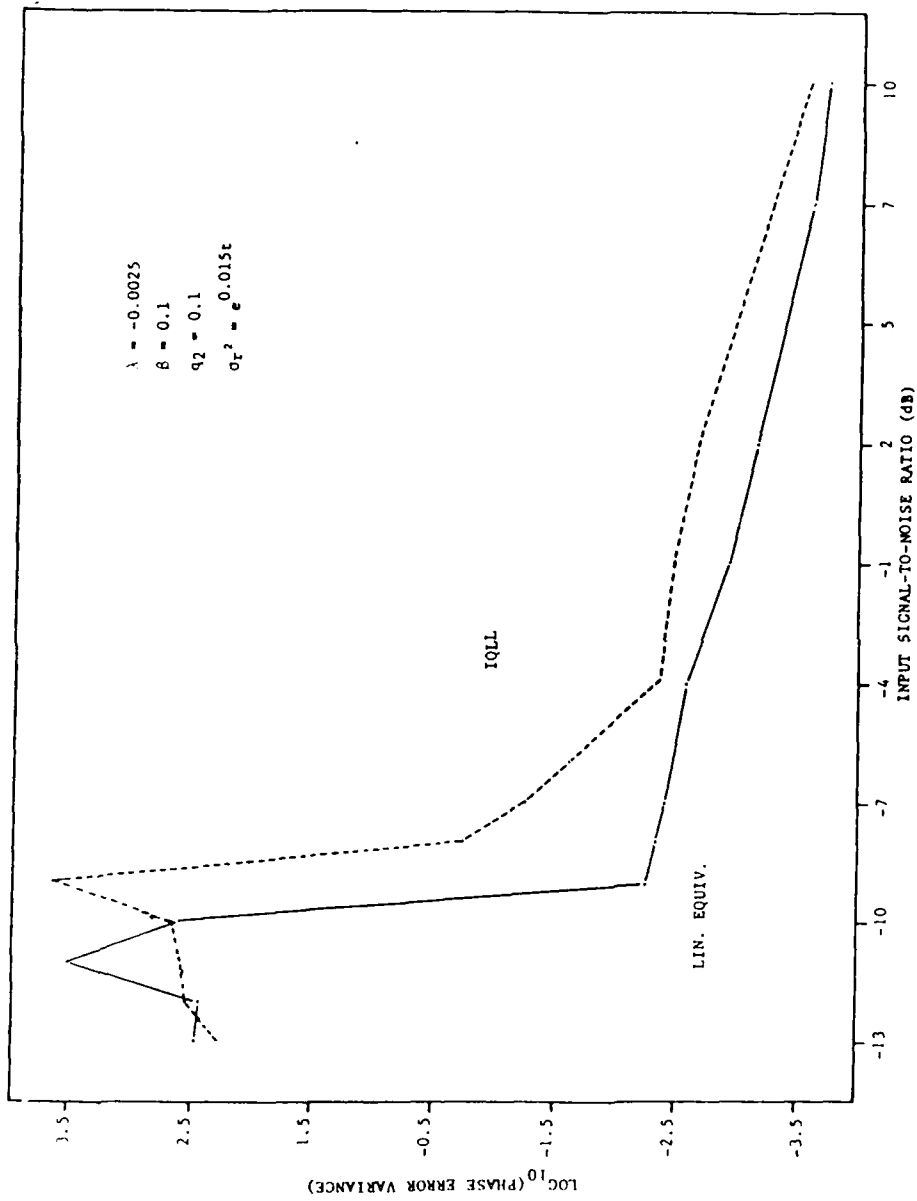
Figure 3-3  MS Phase Estimation Error Comparison; Linear Equivalent versus IQLL
(Advancing target simulation)

IQLL

LIN. EQUIV.

$\lambda = -0.01$
$\beta = 0.1$
$q_2 = 0.01$
$\sigma_{r_0}^2 = 1.0$

INPUT SIGNAL-TO-NOISE RATIO (dB)

$LOG_{10}$(PHASE ERROR VARIANCE)

Figure 3-4   MS Phase Estimation Error Comparison; Linear Equivalent versus IQLL
(Small phase uncertainty simulation)

Figure 3-5  30° RMS Phase Error Cut-off Threshold Comparison; Linear Equivalent
versus IQLL ($\lambda$ = -0.01, $\beta$ = 0.1): Input SNR versus $q_2$.

## TABLE 3-1

### (a) Estimator Sensitivity to Initial Amplitude Variable Uncertainty

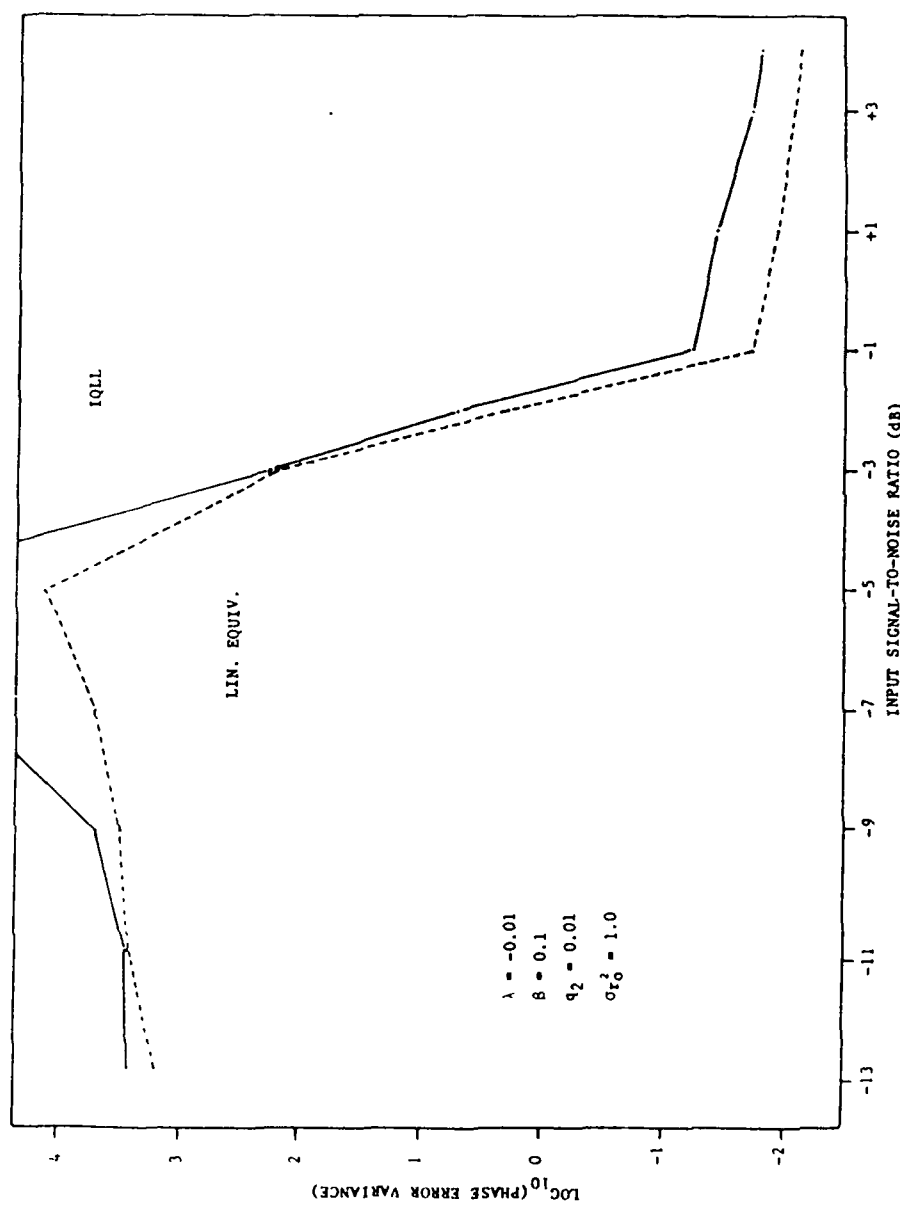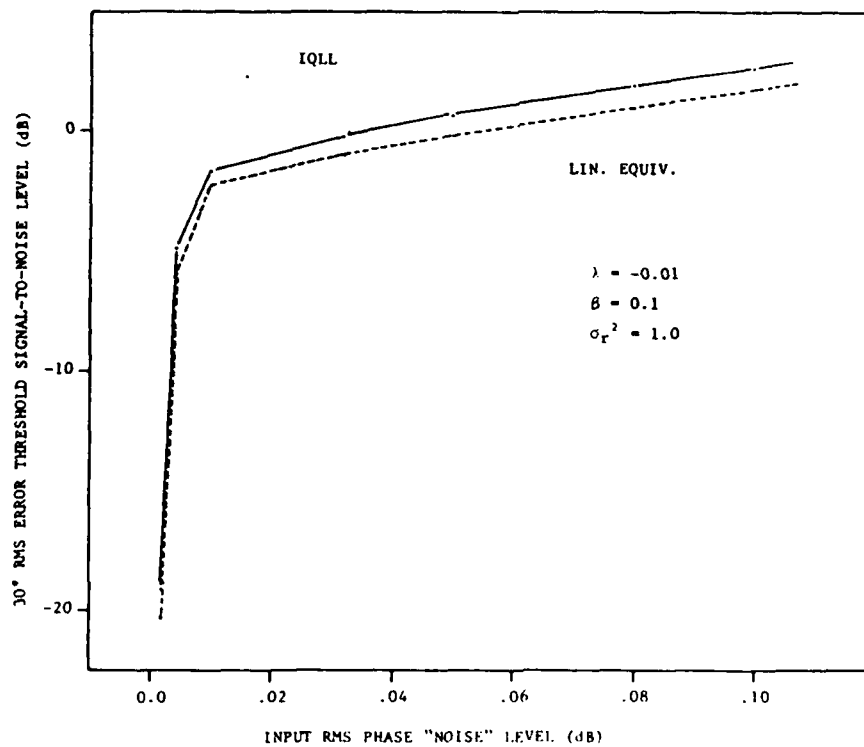| $\sigma^2_{r_o}$ | IQLL | | | | LIN. EQUIV. | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ |
| 160 | 600 | < 300 | < 300 | < 300 | 2400 | $\infty$ | $\infty$ | < 300 |
| 80 | < 300 | < 300 | < 300 | < 300 | 2400 | $\infty$ | $\infty$ | < 300 |
| 45 | < 300 | < 300 | < 300 | < 300 | < 1800 | $\infty$ | $\infty$ | < 300 |
| 40 | < 300 | < 300 | < 300 | < 300 | < 600 | < 300 | < 300 | < 200 |
| 30 | < 100 | < 100 | < 100 | < 100 | < 300 | < 300 | < 300 | < 100 |
| 25 | < 100 | < 100 | < 100 | < 100 | < 100 | < 100 | < 100 | < 100 |

$$(\delta^2_{\theta_o} = \delta^2_{\omega_o} = \delta^2_{\xi_o} = 0)$$

### (b) Estimator Sensitivity to Initial Phase Variable Uncertainty

| $\sigma^2_{\theta_o}$ | IQLL | | | | LIN. EQUIV. | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ |
| 130 | 400 | $\infty$ | 300 | 0 | 400 | $\infty$ | 300 | 0 |
| 110 | 400 | $\infty$ | 300 | 0 | 400 | 600 | 300 | 0 |
| 70 | 300 | $\infty$ | 300 | 0 | 0 | < 400 | 300 | 0 |
| 50 | < 300 | $\infty$ | < 300 | 0 | 0 | 300 | < 300 | 0 |
| 10 | 0 | $\infty$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | $\infty$ | 0 | 0 | 0 | 0 | 0 | 0 |
| 1.5 | 0 | 1500 | 0 | 0 | 0 | 0 | 0 | 0 |
| .75 | 0 | 600 | 0 | 0 | 0 | 0 | 0 | 0 |

(c)  Estimator Sensitivity to Initial Frequency Variable Uncertainty

| $\sigma^2_{\omega_0}$ | IQLL | | | | LIN. EQUIV. | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ |
| 35.0 | 1500 | $\infty$ | $\infty$ | 0 | 1500 | $\infty$ | $\infty$ | 0 |
| 8.0 | 900 | $\infty$ | 1200 | 0 | 900 | $\infty$ | 600 | 0 |
| 4.0 | 600 | $\infty$ | 400 | 0 | 600 | $\infty$ | 600 | 0 |
| 2.0 | 300 | 600 | 300 | < 300 | 300 | 600 | 300 | < 300 |

(d)  Estimator Sensitivity to Initial Frequency Rate Variable Uncertainty

| $\sigma^2_{\xi_0}$ | IQLL | | | | LIN. EQUIV. | | | |
|---|---|---|---|---|---|---|---|---|
| | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ | $T_r$ | $T_\theta$ | $T_\omega$ | $T_\xi$ |
| .05 | 1800 | $\infty$ | $\infty$ | 2400 | 1800 | $\infty$ | $\infty$ | 2400 |
| .025 | 1500 | $\infty$ | 1800 | 1800 | 1800 | $\infty$ | $\infty$ | 1800 |
| .01 | 300 | 900 | 600 | 600 | 300 | 900 | 600 | 600 |

CONVERGENCE ELLIPSOID OUTER BOUNDARY

| INITIAL VARIANCE | | IQLL | LIN. EQUIV. |
|---|---|---|---|
| $\sigma^2_{r_0}$ | < | 220.0 | 45.0 |
| $\sigma^2_{\theta_0}$ | < | 4.0 | 130. |
| $\sigma^2_{\omega_0}$ | < | 3.0 | 2.5 |
| $\sigma^2_{\xi_0}$ | < | .025 | .025 |

TABLE 3-2

## 5. Conclusion

An alternative approach to "amplitude" and phase tracking from quadrature observations has been presented and experimentally investigated. The motivation for the new filtering algorithm comes from the notion of "closer to linear" evolution equations based on linear equivalent dynamics using a Transformation theory for nonlinear diffusion models. Discrete-time approximate filtering equations were formulated for a novel envelope/phase narrowband signal model which should prove useful in real tracking applications in passive sonar. Through Monte Carlo analysis it was shown that the linear equivalent formulation outperforms the IQ locked loop algorithm for estimating "amplitude" and phase in additive noise: as much as 3dB (envelope) and 4-6dB (phase) improvement were demonstrated. This comparison, however, required the augmentation of the standard IQLL algorithm with a fifth state variable (i.e., phase). In fairness to the IQLL approach, which is specifically formulated to avoid this direct estimation of phase, estimates of the in-phase and quadrature coordinate variables were also compared. It was found that the linear equivalent algorithm provided as good or better (up to 1.5dB better) estimates even for these variables as did the IQLL algorithm. Finally, the IQLL appears to be (overall) somewhat less sensitive, in its natural coordinates, than the linear equivalent to accuracy of assumed initial estimates.

Whether the application in mind is for quasi-coherent sequential detection or phase/frequency tracking, the linear equivalent approach appears to offer distinct advantages in terms of mean square error performance. This confirms our original conjecture about the dominant behavior of nonlinear dynamics in the higher SNR regime. On the other hand, when dynamics modelling errors may be significant the IQLL is the preferred approach as the linear equivalent is rather sensitive to the underlying model parameters through the Kalman gain.

## APPENDIX 3.1

### The Linear Representation for the IQ Locked Loop Dynamics

The application of Corollary 2, Chapter 2 to Eqn. (3.5) is a matter of computation, as shown below

Letting $I_t = r_t \cos \theta_t$, $Q_t = r_t \sin \theta_t$, equation (3.5) is reiterated as

$$d \begin{bmatrix} I_t \\ Q_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} \alpha I_t - \omega Q_t \\ \alpha Q_t + \omega I_t \\ \xi_t \\ 0 \end{bmatrix} dt + \begin{bmatrix} \beta I_t & 0 \\ \beta Q_t & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix} \qquad (A-1)$$

Note that this is just the example of Chapter 2, Section 4. Following that example

$$U = \begin{bmatrix} \lambda I & - & \omega Q \\ \lambda Q & + & \omega I \\ & \xi & \\ & 0 & \end{bmatrix}$$

$$g_1 = \text{col} (I, Q, 0, 0)$$

$$g_2 = q_2 \text{ col} (0, 0, 0, 1)$$

$$[U, g_1] = 0 \Rightarrow L_u^m(g_1) = 0 \qquad \forall \; m > 1$$

$$[U, g_2] = \text{col} \; (0, \; 0, \; -q_2, \; 0)$$

$$L_u^2(g_2) = q_2 \; \text{col} \; (-Q, \; I, \; 0, \; 0)$$

$$L_u^3(g_2) = 0$$

The Kronecker indices are $k_1 = 1$, $k_2 = 3$ and

$$L_2(x) = \begin{bmatrix} \beta I & 0 & 0 & -q_2 Q \\ \beta Q & 0 & 0 & q_2 I \\ 0 & 0 & -q_2 & 0 \\ 0 & q_2 & 0 & 0 \end{bmatrix}$$

$\det (L_2(x)) = -\beta q_2^3 \; (I^2 + Q^2) = -\beta q_2^3 \; r^2 \neq 0$ for any $|\beta r| > 0$ and $q_2 \neq 0$. The dual 1-forms are

$$\omega_L(x) = \begin{bmatrix} \dfrac{I}{\beta r^2} & \dfrac{Q}{\beta r^2} & 0 & 0 \\[2mm] 0 & 0 & 0 & 1/q_2 \\[2mm] 0 & 0 & -1/q_2 & 0 \\[2mm] \dfrac{-Q}{q_2 r^2} & \dfrac{I}{q_2 r^2} & 0 & 0 \end{bmatrix} \quad \underline{dx}$$

and $\quad \underline{d}\omega_L = 0 \; \forall \; I, \; q \Rightarrow \omega_L(x) = \underline{d}\theta_L(x)$

i.e., $\omega_L$ may be integrated up exactly, $\theta_L(x) = \int^x \omega_L(\xi)$, and one finds

$$\phi(x) = \theta_L(x) = \begin{bmatrix} \dfrac{1}{2} \ln \; (I^2 + Q^2) \\[2mm] \text{Tan}^{-1}(Q/I) \\[2mm] \omega \\[2mm] \xi \end{bmatrix} \quad \triangleq \quad \begin{bmatrix} \ln(r) \\[2mm] \theta \\[2mm] \omega \\[2mm] \xi \end{bmatrix}$$

Thus the equivalent linear system is

$$
d \begin{bmatrix} Z'_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} \lambda \\ \omega_t \\ \xi_t \\ 0 \end{bmatrix} dt + \begin{bmatrix} \beta & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix}
$$

Letting $Z = Z' - \lambda t$ gives equation (3.6),

$$
d \begin{bmatrix} Z_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Z_t \\ \theta_t \\ \omega_t \\ \xi_t \end{bmatrix} dt + \begin{bmatrix} \beta & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & q_2 \end{bmatrix} \begin{bmatrix} dv_{1t} \\ dv_{2t} \end{bmatrix} \qquad (A-2)
$$

In this case the observation equation is now of the form:

$$
dy_t = \begin{bmatrix} I_t \\ Q_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} = r_o \begin{bmatrix} e^{z_t + \lambda t} \cos\theta_t \\ e^{z_t + \lambda t} \sin\theta_t \end{bmatrix} dt + s \begin{bmatrix} dw_{1t} \\ dw_{2t} \end{bmatrix} \qquad (A-3)
$$

which is equation (7). Of course, the same result is obtained by recognizing that

$\{r_t, \ t \geq 0\}$ is a "log normal" process related to a scaled, time shifted Brownian

motion. However, the equivalence transformation theory provides a systematic method

in the general case. In this sense, it is a tool of some value.

(*) Note here that this transformation results only in a local equivalence,
$\{\theta \mid |\theta| \ll \frac{\pi}{2}\}$. This is not a difficulty, coming only from the choice of the
(I,Q) coordinates to start with. By beginning the application of Corollary 2
with the system in (3.3) one finds a globally equivalent linear diffusion of
the same form as in (A-2).

Chapter 3 - References

[1] K.R. Brown, A. O. Cohen, E. F. Harrold, & G.W. Johnson, IQ-Locked Loop tracker, ICASSP, 1976

[2] G.W. Johnson, K.R. Brown , A.O Cohen & E.F. Harrold, Improved frequency line tracking, Interim Report IRAD Proj. #SM70, IBM/FSD Manassas, Va. (Report #75-0101-M19) 17 July 1975

[3] S.C. Schwartz & Y.S. Yeh, On the distribution function & moments of Power Sums with Log Normal Components, Bell System Tech. Jour., Vol. 61, No. 7 Sept., 1982, pp 1441-1462

[4] L. Fahrmeir, On the simulation of Stochastic Processes with Continuous State Space and Parameter Space, in Simulation of Systems, L. Dekker, edit., North-Holland 1976, pp 67-71

[5] H. Flanders, Differential Forms with Application to the Physical Sciences, Academic Press, 1963

[6] N. Ikeda & S. Watanabe, Stochastic Differential Equations and Diffusions Processes, North Holland, 1981 (especially Chap. V)

OFFICE OF NAVAL RESEARCH

Report Distribution List
Contract N00014-84-C-0233

COPIES

Head, Mathematical Sciences Division
Office of Naval Research
Arlington, VA  22217                                    1


Defense Technical Information Center
Bldg. 5, Cameron Station
Alexandria, Virginia  22314                            12


Administrative Contracting Officer
DCASMA
P.O. Box 7699
Philadelphia, PA  19101                                 1


Director
Naval Research Laboratory
Attn:  Code 2627
Washington, DC  20375                                   6


Mr. James Smith, Code NAIR 340R
Naval Air Systems Command
Washington, DC  20361                                   1

# END

# FILMED

7-85

# DTIC